REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, ir~~cluding the time for reviewing instructions, searching existing data so~~urces, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send ~~comments regarding this burden estimate or any other aspect of t~~his collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Dir~~ectorate for Information Operations and Reports, 1215 Jefferson Davi~~s Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of management and Budget, Paperwork Reduction P~~roject (0704-0188), Washington, DC 20503.~~

AFRL-SR-BL-TR-98-

0796

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE December, 1994	3. Final

4. TITLE AND SUBTITLE USAF Summer Research Program - 1994 Summer Faculty Research Program Final Reports, Volume 3B, Phillips Laboratory	5. FUNDING NUMBERS
6. AUTHORS Gary Moore	

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Research and Development Labs, Culver City, CA	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFOSR/NI 4040 Fairfax Dr, Suite 500 Arlington, VA 22203-1613	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Contract Number: F49620-93-C-0063

12a. DISTRIBUTION AVAILABILITY STATEMENT Approved for Public Release	12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The United States Air Force Summer Faculty Research Program (USAF- SFRP) is designed to introduce university, college, and technical institute faculty members to Air Force research. This is accomplished by the faculty members being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory Technical Directorates and Air Force Air Logistics Centers. Each participant provided a report of their research, and these reports are consolidated into this annual report.

14. SUBJECT TERMS AIR FORCE RESEARCH, AIR FORCE, ENGINEERING, LABORATORIES, REPORTS, SUMMER, UNIVERSITIES	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1994

SUMMER FACULTY RESEARCH PROGRAM FINAL REPORTS

VOLUME 3B

PHILLIPS LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major David Hart

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Gwendolyn Smith

Program Administrator
Johnetta Thompson

Submitted to:

DTIC QUALITY INSPECTED 4

PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.

Due to its length, Volume 3 is bound in two parts, 3A and 3B. Volume 3A contains #1-19. Volume 3B contains reports #20-39. The Table of Contents for Volume 3 is included in both parts.

This document is one of a set of 16 volumes describing the 1994 AFOSR Summer Research Program. The following volumes comprise the set:

VOLUME	TITLE
1	Program Management Report
	Summer Faculty Research Program (SFRP) Reports
2A & 2B	Armstrong Laboratory
3A & 3B	Phillips Laboratory
4	Rome Laboratory
5A & 5B	Wright Laboratory
6	Arnold Engineering Development Center, Frank J. Seiler Research Laboratory, and Wilford Hall Medical Center
	Graduate Student Research Program (GSRP) Reports
7	Armstrong Laboratory
8	Phillips Laboratory
9	Rome Laboratory
10	Wright Laboratory
11	Arnold Engineering Development Center, Frank J. Seiler Research Laboratory, and Wilford Hall Medical Center
	High School Apprenticeship Program (HSAP) Reports
12A & 12B	Armstrong Laboratory
13	Phillips Laboratory
14	Rome Laboratory
15A&15B	Wright Laboratory
16	Arnold Engineering Development Center

SFRP FINAL REPORT TABLE OF CONTENTS   **i-xxi**

SFRP FINAL REPORTS

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

vii

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

xiii

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

SRP Final Report Table of Contents

1. INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

The numbers of projected summer research participants in each of the three categories are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest; and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1994 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the 1990 contract, RDL won the recompetition for the basic year and four 1-year options.

1

2.   PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

Table 1: SRP Participation, by Year

YEAR	Number of Participants			TOTAL
	SFRP	GSRP	HSAP	
1979	70			70
1980	87			87
1981	87			87
1982	91	17		108
1983	101	53		154
1984	152	84		236
1985	154	92		246
1986	158	100	42	300
1987	159	101	73	333
1988	153	107	101	361
1989	168	102	103	373
1990	165	121	132	418
1991	170	142	132	444
1992	185	121	159	464
1993	187	117	136	440
1994	192	117	133	442

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years; in one case a laboratory did not fund any additional associates. However, the table shows that, overall, the number of participating associates increased this year because two laboratories funded more associates than they had in previous years.

3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 44-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous (over 600 annually) individual requesters.

Due to a delay in awarding the new contract, RDL was not able to place advertisements in any of the following publications in which the SRP is normally advertised: *Black Issues in Higher Education, Chemical & Engineering News, IEEE Spectrum* and *Physics Today*.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1994 are shown in this table.

Table 2: 1994 Applicants and Participants

PARTICIPANT CATEGORY	TOTAL APPLICANTS	SELECTEES	DECLINING SELECTEES
SFRP	600	192	30
(HBCU/MI)	(90)	(16)	(7)
GSRP	322	117	11
(HBCU/MI)	(11)	(6)	(0)
HSAP	562	133	14
TOTAL	1484	442	55

4. SITE VISITS

During June and July of 1994, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

In previous years, an RDL program representative visited from seven to ten different HBCU/MIs to promote interest in the SRP among the faculty and graduate students. Due to the late contract award date (January 1994) no time was available to visit HBCU/MIs this past year.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, seven HBCU/MI SFRPs declined after they were selected. The following table records HBCU/MI participation in this program.

Table 3: SRP HBCU/MI Participation, by Year

YEAR	SFRP		GSRP	
	Applicants	Participants	Applicants	Participants
1985	76	23	15	11
1986	70	18	20	10
1987	82	32	32	10
1988	53	17	23	14
1989	39	15	13	4
1990	43	14	17	3
1991	42	13	8	5
1992	70	13	9	5
1993	60	13	6	2
1994	90	16	11	6

6. SRP FUNDING SOURCES

Funding sources for the 1994 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1994 SRP selected participants are shown here.

Table 4: 1994 SRP Associate Funding

FUNDING CATEGORY	SFRP	GSRP	HSAP
AFOSR Basic Allocation Funds	150	98[*1]	121[*2]
USAF Laboratory Funds	37	19	12
HBCU/MI By AFOSR (Using Procured Addn'l Funds)	5	0	0
TOTAL	192	117	133

*1 - 100 were selected, but two canceled too late to be replaced.
*2 - 125 were selected, but four canceled too late to be replaced.

7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

Table 5: 1994 SRP Associate Compensation

PARTICIPANT CATEGORY	1991	1992	1993	1994
Faculty Members	$690	$718	$740	$740
Graduate Student (Master's Degree)	$425	$442	$455	$455
Graduate Student (Bachelor's Degree)	$365	$380	$391	$391
High School Student (First Year)	$200	$200	$200	$200
High School Student (Subsequent Years)	$240	$240	$240	$240

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $37/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and GSRP associates, 58% (178 out of 309) claimed travel reimbursements at an average round-trip cost of $860.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-one percent (78 out of 192) of faculty associates took orientation trips at an average cost of $498. Many faculty associates noted on their evaluation forms that due to the late notice of acceptance into the 1994 SRP (caused by the late award in January 1994 of the contract) there wasn't enough time to attend an orientation visit prior to their tour start date. In 1993, 58 % of SFRP associates took orientation visits at an average cost of $685.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

8.   CONTENTS OF THE 1994 REPORT

The complete set of reports for the 1994 SRP includes this program management report augmented by fifteen volumes of final research reports by the 1994 associates as indicated below:

Table 6: 1994 SRP Final Report Volume Assignments

LABORATORY	VOLUME		
	SFRP	GSRP	HSAP
Armstrong	2	7	12
Phillips	3	8	13
Rome	4	9	14
Wright	5A, 5B	10	15
AEDC, FJSRL, WHMC	6	11	16

AEDC   =   Arnold Engineering Development Center
FJSRL   =   Frank J. Seiler Research Laboratory
WHMC   =   Wilford Hall Medical Center

APPENDIX A -- PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP and GSRP associates represent 158 different colleges, universities, and institutions.

B. States Represented

SFRP -Applicants came from 46 states plus Washington D.C. and Puerto Rico. Selectees represent 40 states.

GSRP - Applicants came from 46 states and Puerto Rico. Selectees represent 34 states.

HSAP - Applicants came from fifteen states. Selectees represent ten states.

C. Academic Disciplines Represented

The academic disciplines of the combined 192 SFRP associates are as follows:

Electrical Engineering	22.4%
Mechanical Engineering	14.0%
Physics: General, Nuclear & Plasma	12.2%
Chemistry & Chemical Engineering	11.2%
Mathematics & Statistics	8.1%
Psychology	7.0%
Computer Science	6.4%
Aerospace & Aeronautical Engineering	4.8%
Engineering Science	2.7%
Biology & Inorganic Chemistry	2.2%
Physics: Electro-Optics & Photonics	2.2%
Communication	1.6%
Industrial & Civil Engineering	1.6%
Physiology	1.1%
Polymer Science	1.1%
Education	0.5%
Pharmaceutics	0.5%
Veterinary Medicine	0.5%
TOTAL	100%

Table A-1. Total Participants

Number of Participants	
SFRP	192
GSRP	117
HSAP	133
TOTAL	442

Table A-2. Degrees Represented

Degrees Represented			
	SFRP	GSRP	TOTAL
Doctoral	189	0	189
Master's	3	47	50
Bachelor's	0	70	70
TOTAL	192	117	309

Table A-3. SFRP Academic Titles

Academic Titles	
Assistant Professor	74
Associate Professor	63
Professor	44
Instructor	5
Chairman	1
Visiting Professor	1
Visiting Assoc. Prof.	1
Research Associate	3
TOTAL	192

Table A-4. Source of Learning About SRP

SOURCE	SFRP		GSRP	
	Applicants	Selectees	Applicants	Selectees
Applied/participated in prior years	26%	37%	10%	13%
Colleague familiar with SRP	19%	17%	12%	12%
Brochure mailed to institution	32%	18%	19%	12%
Contact with Air Force laboratory	15%	24%	9%	12%
Faculty Advisor (GSRPs Only)	--	--	39%	43%
Other source	8%	4%	11%	8%
TOTAL	100%	100%	100%	100%

Table A-5. Ethnic Background of Applicants and Selectees

	SFRP		GSRP		HSAP	
	Applicants	Selectees	Applicants	Selectees	Applicants	Selectees
American Indian or Native Alaskan	0.2%	0%	1%	0%	0.4%	0%
Asian/Pacific Islander	30%	20%	6%	8%	7%	10%
Black	4%	1.5%	3%	3%	7%	2%
Hispanic	3%	1.9%	4%	4.5%	11%	8%
Caucasian	51%	63%	77%	77%	70%	75%
Preferred not to answer	12%	14%	9%	7%	4%	5%
TOTAL	100%	100%	100%	100%	99%	100%

Table A-6. Percentages of Selectees receiving their 1st, 2nd, or 3rd Choices of Directorate

	1st Choice	2nd Choice	3rd Choice	Other Than Their Choice
SFRP	70%	7%	3%	20%
GSRP	76%	2%	2%	20%

APPENDIX B -- SRP EVALUATION RESPONSES

1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

Evaluation Group	Responses
SFRP & GSRPs	275
HSAPs	116
USAF Laboratory Focal Points	109
USAF Laboratory HSAP Mentors	54

All groups indicate near-unanimous enthusiasm for the SRP experience.

Typical comments from 1994 SRP associates are:

"[The SRP was an] excellent opportunity to work in state-of-the-art facility with top-notch people."

"[The SRP experience] enabled exposure to interesting scientific application problems; enhancement of knowledge and insight into 'real-world' problems."

"[The SRP] was a great opportunity for resourceful and independent faculty [members] from small colleges to obtain research credentials."

"The laboratory personnel I worked with are tremendous, both personally and scientifically. I cannot emphasize how wonderful they are."

"The one-on-one relationship with my mentor and the hands on research experience improved [my] understanding of physics in addition to improving my library research skills. Very valuable for [both] college and career!"

Typical comments from laboratory focal points and mentors are:

"This program [AFOSR - SFRP] has been a 'God Send' for us. Ties established with summer faculty have proven invaluable."

"Program was excellent from our perspective. So much was accomplished that new options became viable "

"This program managed to get around most of the red tape and 'BS' associated with most Air Force programs. Good Job!"

'Great program for high school students to be introduced to the research environment. Highly educational for others [at laboratory]."

'This is an excellent program to introduce students to technology and give them a feel for [science/engineering] career fields. I view any return benefit to the government to be 'icing on the cake' and have usually benefitted."

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below (Note: basically the same as in previous years.)

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space. computer assets, clearly defined scope of work).

B. Laboratory sponsor seminar presentations of work conducted by associates, and/or organized social functions for associates to collectively meet and share SRP experiences.

C. Laboratory focal points collectively suggest more AFOSR allocated associate positions. so that more people may share in the experience.

D. Associates collectively suggest higher stipends for SRP associates.

E. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same arguement was used to raise the faculty and graduate student participation time a few years ago.)

2. 1994 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 109 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

| | | How Many Associates Would You Prefer To Get ? | | | | | | | | (% Response) | | | |
| | | SFRP | | | | GSRP (w/Univ Professor) | | | | GSRP (w/o Univ Professor) | | | |
Lab	Evals Recv'd	0	1	2	3+	0	1	2	3+	0	1	2	3+
AEDC	10	30	50	0	20	50	40	0	10	40	60	0	0
AL	44	34	50	6	9	54	34	12	0	56	31	12	0
FJSRL	3	33	33	33	0	67	33	0	0	33	67	0	0
PL	14	28	43	28	0	57	21	21	0	71	28	0	0
RL	3	33	67	0	0	67	0	33	0	100	0	0	0
WHMC	1	0	0	100	0	0	100	0	0	0	100	0	0
WL	46	15	61	24	0	56	30	13	0	76	17	6	0
Total	121	25%	43%	27%	4%	50%	37%	11%	1%	54%	43%	3%	0%

LFP Evaluation Summary. The summarized repsonses. by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
    a. Time available for evaluation of applications:
    b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5   a. Benefits of associate's work to laboratory:
    b. Benefits of associate's work to Air Force:
6.   a. Enhancement of research qualifications for LFP and staff:
    b. Enhancement of research qualifications for SFRP associate:
    c. Enhancement of research qualifications for GSRP associate:
7.   a. Enhancement of knowledge for LFP and staff:
    b. Enhancement of knowledge for SFRP associate:
    c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10.   a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
     b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Laboratory Focal Point Reponses to above questions

	AEDC	AL	FJSRL	PL	RL	WHMC	WL
= Evals Recv'd	10	32	3	14	3	1	46
Question #							
2	90 %	62 %	100 %	64 %	100 %	100 %	83 %
2a	**3.5**	**3.5**	4.7	4.4	4.0	4.0	**3.7**
2b	4.0	3.8	4.0	4.3	4.3	4.0	3.9
3	4.2	3.6	4.3	3.8	4.7	4.0	4.0
4	3.8	3.9	4.0	4.2	4.3	NO ENTRY	4.0
5a	4.1	4.4	4.7	4.9	4.3	3.0	4.6
5b	4.0	4.2	4.7	4.7	4.3	3.0	4.5
6a	**3.6**	4.1	**3.7**	4.5	4.3	**3.0**	4.1
6b	3.6	4.0	4.0	4.4	4.7	3.0	4.2
6c	3.3	4.2	4.0	4.5	4.5	3.0	4.2
7a	3.9	4.3	4.0	4.6	4.0	3.0	4.2
7b	4.1	4.3	4.3	4.6	4.7	3.0	4.3
7c	3.3	4.1	4.5	4.5	4.5	5.0	4.3
8	4.2	4.3	5.0	4.9	4.3	5.0	4.7
9	3.8	4.1	4.7	5.0	4.7	5.0	4.6
10a	4.6	4.5	5.0	4.9	4.7	5.0	4.7
10b	4.3	4.2	5.0	4.3	5.0	5.0	4.5
11	4.1	4.5	4.3	4.9	4.7	4.0	4.4
12	4.1	3.9	4.0	4.4	4.7	3.0	4.1
13a	**3.8**	**2.9**	4.0	4.0	4.7	**3.0**	**3.6**
13b	**3.8**	**2.9**	4.0	4.3	4.7	**3.0**	**3.8**
14	4.5	4.4	5.0	4.9	4.7	4.0	4.5

3. 1994 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 275 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. The match between the laboratories research and your field:	4.6
2. Your working relationship with your LFP:	4.8
3. Enhancement of your academic qualifications:	4.4
4. Enhancement of your research qualifications:	4.5
5. Lab readiness for you: LFP, task, plan:	4.3
6. Lab readiness for you: equipment, supplies, facilities:	4.1
7. Lab resources:	4.3
8. Lab research and administrative support:	4.5
9. Adequacy of brochure and associate handbook:	4.3
10. RDL communications with you:	4.3
11. Overall payment procedures:	**3.8**
12. Overall assessment of the SRP:	4.7

13. a. Would you apply again?	Yes:	85%
b. Will you continue this or related research?	Yes:	95%

14. Was length of your tour satisfactory?	Yes:	86%

15. Percentage of associates who engaged in:

a. Seminar presentation:	52%
b. Technical meetings:	32%
c. Social functions:	03%
d. Other	01%

16. Percentage of associates who experienced difficulties in:

    a. Finding housing: 12%
    b. Check Cashing: 03%

17. Where did you stay during your SRP tour?

    a. At Home: 20%
    b. With Friend: 06%
    c. On Local Economy: 47%
    d. Base Quarters: 10%

THIS SECTION FACULTY ONLY:

18. Were graduate students working with you?   Yes: 23%

19. Would you bring graduate students next year?   Yes: 56%

20. Value of orientation visit:

    Essential: 29%
    Convenient: 20%
    Not Worth Cost: 01%
    Not Used: 34%

THIS SECTION GRADUATE STUDENTS ONLY:

21. Who did you work with:

    University Professor: 18%
    Laboratory Scientist: 54%

4. 1994 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

The summarized results listed below are from the 54 mentor evaluations received.

1. Mentor apprentice preferences:

Table B-3. Air Force Mentor Responses

Laboratory	# Evals Recv'd	How Many Apprentices Would You Prefer To Get ?			
		HSAP Apprentices Preferred			
		0	1	2	3+
AEDC	6	0	100	0	0
AL	17	29	47	6	18
PL	9	22	78	0	0
RL	4	25	75	0	0
WL	18	22	55	17	6
Total	54	20%	71%	5%	5%

Mentors were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

2. Mentors involved in SRP apprentice application evaluation process:
    a. Time available for evaluation of applications:
    b. Adequacy of applications for selection process:
3. Laboratory's preparation for apprentice:
4. Mentor's preparation for apprentice:
5. Length of research tour:
6. Benefits of apprentice's work to U.S. Air force:
7. Enhancement of academic qualifications for apprentice:
8. Enhancement of research skills for apprentice:
9. Value of U.S. Air Force/high school links:
10. Mentor's working relationship with apprentice:
11. Expenditure of mentor's time worthwhile:
12. Quality of program literature for apprentice:
13.   a. Quality of RDL's communications with mentors:
    b. Quality of RDL's communication with apprentices:
14. Overall assessment of SRP:

	AEDC	AL	PL	RL	WL
# Evals Recv'd	6	17	9	4	18
Question #					
2	100 %	76 %	56 %	75 %	61 %
2a	4.2	4.0	**3.1**	**3.7**	**3.5**
2b	4.0	4.5	4.0	4.0	3.8
3	4.3	**3.8**	**3.9**	**3.8**	**3.8**
4	4.5	**3.7**	**3.4**	4.2	**3.9**
5	**3.5**	4.1	**3.1**	**3.7**	**3.6**
6	4.3	3.9	4.0	4.0	4.2
7	4.0	4.4	4.3	4.2	3.9
8	4.7	4.4	4.4	4.2	4.0
9	4.7	4.2	3.7	4.5	4.0
10	4.7	4.5	4.4	4.5	4.2
11	4.8	4.3	4.0	4.5	4.1
12	4.2	4.1	4.1	4.8	3.4
13a	**3.5**	**3.9**	**3.7**	4.0	**3.1**
13b	4.0	4.1	3.4	4.0	3.5
14	4.3	4.5	3.8	4.5	4.1

5. 1994 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 116 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from
1 (below average) to 5 (above average)

1. Match of lab research to you interest:	3.9
2. Apprentices working relationship with their mentor and other lab scientists:	4.6
3. Enhancement of your academic qualifications:	4.4
4. Enhancement of your research qualifications:	4.1
5. Lab readiness for you: mentor. task. work plan	**3.7**
6. Lab readiness for you: equipment supplies facilities	4.3
7. Lab resources: availability	4.3
8. Lab research and administrative support:	4.4
9. Adequacy of RDL's apprentice handbook and administrative materials:	4.0
10. Responsiveness of RDL's communications:	**3.5**
11. Overall payment procedures:	**3.3**
12. Overall assessment of SRP value to you:	4.5

13. Would you apply again next year? Yes: 88%

14. Was length of SRP tour satisfactory? Yes: 78%

15. Percentages of apprentices who engaged in:

a. Seminar presentation:	48%
b. Technical meetings:	23%
c. Social functions:	18%

INVESTIGATIONS OF ELECTRON INTERACTIONS WITH MOLECULES: ELECTRON ATTACHMENT AND ELECTRON DETACHMENT REACTIONS OF HALOGENATED MOLECULES

Albert D. Kowalak
Professor
Department of Chemistry
University of Massachusetts Lowell
1 University Ave
Lowell, MA 01854

and

Jane M. Van Doren
Assistant Professor
Department of Chemistry
College of the Holy Cross
1 College Street
Worcester, MA 01610-2395

INVESTIGATIONS OF ELECTRON INTERACTIONS WITH MOLECULES: ELECTRON ATTACHMENT AND ELECTRON DETACHMENT REACTIONS OF HALOGENATED MOLECULES

Jane M. Van Doren   Albert D. Kowalak

ABSTRACT

Reactions of electrons with a variety of halogenated molecules were explored using a Flowing Afterglow Langmuir Probe and mass spectrometer. Electron attachment to octafluorotoluene, pentafluorobenzonitrile, pentafluoroacetophenone, trifluorotolunitrile, and pentafluoropyridine results in the formation of a parent anion only, i.e., the association product is formed. Rate coefficients for electron attachment to these molecules indicate that these reactions are generally efficient (at 300K, k(octafluorotoluene) = 2.1 x 10 cm^3 s^{-1}, k(pentafluorobenzonitrile) $\geq$ 2.3 x 10 cm^3 s^{-1}, k(pentafluoropyridine) = 1.8 x 10 cm^3 s^{-1}). In two cases, reactions of octafluorotoluene and pentafluoropyridine, the anion formed is observed to detach electrons under our experimental conditions. Analysis of these data assuming that detachment is thermally controlled, indicates that the electron affinities of octafluorotoluene and pentafluoropyridine are < 1 eV. Chloroacetonitrile and pentafluorobenzene react with electrons to form ions which are fragments of the parent, i.e., dissociative attachment products. Attachment to chloroacetonitrile forming Cl^- is relatively efficient, k = 4 x 10 cm^3 s^{-1} at 300K and becomes increasingly more efficient as temperature is increased. Electron attachment to pentafluorobenzene appears to form at least two products arising from attachment followed by loss of an H atom and attachment followed by loss of a HF molecule. Work has begun on an analytical strategy to extend the range of molecules one can study with the Flowing Afterglow Langmuir Probe technique.

INVESTIGATIONS OF ELECTRON INTERACTIONS WITH MOLECULES: ELECTRON ATTACHMENT AND ELECTRON DETACHMENT REACTIONS OF HALOGENATED MOLECULES

Jane M. Van Doren   Albert D. Kowalak

Introduction

The interactions of electrons with molecules play important roles in many systems including electronics, biochemistry and atmospheric chemistry. One important area of research is the investigation of reactions which chemically transform electrons to negative ions. Such reactions play central roles in the production of ions in the atmosphere, determination of the atmospheric lifetime of neutral (reactant) molecules and the evolution of natural and man-made plasmas. The Air Force research program has contributed to our understanding of many of these processes. In recent years, the Air Force emphasis has been on plasma evolution and to a lessor extent, atmospheric lifetimes of trace species.

Information on interactions between electrons and molecules is relatively scarce, reflecting the difficulty in both experimentally probing electron-molecule interactions and interpreting the results of such probes. The Flowing Afterglow Langmuir Probe (FALP) Technique is a versatile technique which allows investigation of a wide range of electron-molecule reactions at low collision energies, providing quantitative values for electron attachment rate coefficients. Coupling the FALP with a mass spectrometer

allows product identification which is especially important for reactions in which more than one reaction pathway is thermochemically allowed or in which the products formed change with temperature. In a small number of reactions, two different product ions are formed. In this case, the mass spectrometer enables one to determine the product branching fraction or relative rate coefficient for each product channel. The experiments reported and discussed in this report were performed with the FALP with mass spectrometer detection available at the Phillips Laboratory. This analytical research instrument is one of only three such systems in the world.

The experiments discussed below focussed on reactions of electrons with a variety of halogenated molecules. We chose to study these systems because data in the literature indicates that introduction of halogen atoms into a molecule can result in the stabilization of the parent negative ion. Species which can form stable anionic products upon addition of an electron can potentially do so efficiently. Our investigations attempt to reveal some of the important chemical features which result in efficient electron attachment and to provide information about the relative stability of the product ions. Finally, in the process of our investigations, new applications of this experimental technique and new analytical strategies were probed.

Experimental

The Flowing Afterglow Langmuir Probe Technique (FALP) utilizes a Langmuir probe which measures electron density. In our experiment, a plasma is created by a microwave discharge in Helium (~1 torr). A small flow of Argon is introduced downstream of the discharge to eliminate the small fraction of helium metastable species formed. The electrons and positive ions in the plasma are carried down the length of the flow tube by a fast laminar flow of buffer gas. The electron density may be measured at any point along the length of the reaction flow tube with a moveable Langmuir Probe. Neutral reactants are added through a radial inlet downstream of the plasma. Ionic products of electron attachment reactions may be identified with a quadrupole mass spectrometer which is located at the downstream end of the flow tube and separated from the relatively high pressure tube by a small orifice. Most of the fast flowing gas is

pumped by a large roots blower and mechanical pumping system. A schematic of the instrument may be found in Figure 1.

A rate coefficient for electron attachment is determined from the relative decrease in electron density with reaction distance (which is proportional to reaction time in this "steady-state" experiment). The observed electron loss in the presence of an attaching gas is determined by diffusive loss to the flow tube walls and reactive loss from reaction with the reactant neutral. These loss rates are coupled. Therefore, loss by ambipolar diffusion is evaluated separately by monitoring the electron density as a function of distance in the absence of an attaching gas. Using this diffusion rate and the observed electron loss in the presence of an attaching gas, one numerically solves the equation as given by Smith and Spanel[1] for the electron attachment rate coefficient.

Results and Discussion

Most of the molecules investigated react with electrons by associative attachment, where the product ion is the anion of the reactant molecule. Other molecules were found to react with electrons by dissociative attachment, where the product ion is a fragment of the reactant molecule. Rate coefficients were determined and product ions identified for most of the reactions studied. The effect of temperature on the rate coefficient and product ions formed was evaluated for several systems. For associative detachment reactions, pressure dependences of the rate coefficient for attachment were evaluated. In some systems, subsequent detachment of electrons from the product ions formed was also observed. Preliminary analysis and discussion of this process is discussed below. Finally, several of the compounds investigated had either low vapor pressures or were highly reactive. These qualities make determination of electron attachment rate coefficients difficult. Following the lead of Smith et al.,[2] we have begun developing an analytical procedure for determining rate coefficients under these conditions.

FLOWING AFTERGLOW LANGMUIR PROBE

Figure 1. Schematic diagram of Flowing Afterglow Langmuir Probe apparatus with mass spectrometer

Associative Attachment Reactions forming stable negative ions at 300K.

Octafluorotoluene, pentafluoropyridine, pentafluorobenzonitrile, pentafluoroacetophenone, trifluorotolunitrile all react with thermal electrons at 300K by associative attachment as evidenced by the mass spectra (see for example, Figure 2), i.e., only the parent negative ion is formed.

Electron attachment to octafluorotoluene and pentafluoropyridine was found to be very efficient with rate coefficients of 2.1×10^{-7} cm^3 s^{-1} and 2.1×10^{-7} cm^3 s^{-1}, respectively. Reaction of electrons with pentafluorobenzonitrile and pentafluoroacetophenone is also efficient. Accurate determination of the rate coefficient for electron attachment for these compounds proved difficult because the measured value appeared to depend on the length of time the reactant had been flowing in the inlet line (distinguished from the time for reaction with electrons in the flow tube). Such behavior, discussed further below , is indicative of species which can and do react with trace species adsorbed on the inlet walls or react with the inlet walls themselves.

An estimate of the electron attachment rate coefficient can be obtained for such systems by monitoring the observed rate coefficient as a function of time the vapor has been flowing in the inlet lines (see for example, Figure 3). In general the flow of the reactant is reduced from the value measured at the flowmeter by reactions with the inlet walls. At long times, reaction will be complete or will reach equilibrium and the neutral reactant flow will be equal to the value measured by the flowmeter. Therefore, the value obtained at long times should provide a lower limit to the true rate coefficient. This explanation appears to be supported by the data obtained for pentafluorobenzonitrile. Using the time dependence of the observed rate coefficient we estimate that the true electron attachment rate coefficient is $\geq 2.3 \times 10^{-7}$ cm^3 s^{-1}. In contrast, the measured rate coefficient for electron attachment to pentafluoroacetophenone did not appear to reach a maximum value over the time measured. At this point we are unsure of the implications of these measurements. We are currently developing an analytical procedure which would allow simultaneous neutral concentration and electron attachment rate coefficient determination (see below.) No rate coefficient was measured for electron attachment to trifluorotolunitrile because of the very low vapor pressure of this compound.

Figure 2. Mass spectrum obtained upon addition of octafluorotoluene to the reaction flow tube.

Figure 3. Observed electron attachment rate coefficient as a function of time the pentafluorobenzonitrile vapor has been flowing in the inlet lines .

Interpretation of the measured electron attachment rate coefficient for reactions forming an association product requires an understanding of the reaction mechanism. The initial association product is highly excited and has enough energy to detach the electron.

$$X + e^- \underset{k_d}{\overset{k_a}{\Leftrightarrow}} X^{-*}$$

$$X^{-*} + He \overset{\alpha k_c}{\Rightarrow} X^-$$

The bimolecular rate coefficient for association is given by k_a while the unimolecular rate coefficient for detachment of the electron from the energetically excited association product is given by k_d. Only those excited association products which undergo an inelastic collision with the surrounding molecules (typically He at pressure of 1 torr) to lose energy and form a stable (bound) negative ion result in reaction, or loss of electrons as measured in our millisecond time scale experiment. Assuming the the concentration of $[X^{-*}]$ is in a steady state, we find

$$d[e^-]/dt = - k_a[X][e] \{1 - k_d/(k_d + \alpha k_c[He])\}$$

In the high pressure limit, the rate of electron loss is independent of the pressure of collision gas He and depends only on the initial association rate coefficient. In this case, the measured rate coefficient would be equal to the association rate coefficient. We have measured the rate coefficient for electron attachment to octafluorotoluene and pentafluoropyridine as a function of flow tube (collision gas) pressure and found the observed attachment rate coefficients are independent of this parameter. Therefore, in the cases of octafluorotoluene and pentafluoropyridine the measured rate coefficients are a direct measure of the association rate coefficients. Similar studies must be performed for the reactions of the other associatively attaching species.

Associative Attachment Reactions followed by Electron Detachment

Weakly bound anions can detach in the reaction flow tube. When anions initially formed by attachment of electrons subsequently detach, the observed loss of electrons due to the presence of an

attaching gas decreases. In our steady-state flow reactor, this appears as a positive deviation from the electron density expected in the presence of a single well-behaved attaching gas. This deviation is most pronounced at longer distances because at those distances a substantial number of anions (which can detach) are present (see for example, Figure 4).

We observed evidence of electron detachment in the reactions of pentafluoropyridine (at temperature $T \geq 465K$) and octafluorotoluene ($T \geq 510K$). The only product ion observed in both reactions is the parent anion, the anion formed from associative attachment. In such cases, we can analyze the electron density data quantitatively to find values for the electron attachment and detachment rate coefficients. Specifically , we fit the observed loss of electrons with distance with values for the attachment and detachment rate coefficients assuming that electrons are lost by attachment and diffusion and "gained" by detachment. It is important to note that the constraints are provided by the stipulation that negative ions do not diffuse and that the sum of electrons plus negative ions must equal the number of positive ions at all times. As in our attachment rate coefficient analysis, the diffusion rate is not a variable but is determined in a separate experiment in which the attaching gas is not present. The recognition that these rate coefficients could be evaluated and the procedure discussed above was recently presented by Miller et al.[3] Electron attachment attachment and detachment rate coefficients for pentafluoropyridine and octafluorotoluene as a function of temperature are presented in Figures 5 and 6.

Interpretation of the observed rate coefficients requires an understanding of the mechanism of detachment. If the observed detachment is truly a thermal process the relative values of the rate coefficients, k_a/k_d, is a measure of the equilibrium constant K as defined by

$$X + e^- \underset{k_d}{\overset{k_a}{\Leftrightarrow}} X^- \quad \text{, where } k_a/k_d = K$$

Such an analysis was presented by Miller et al.[3] This equilibrium constant is related to the exoergicity of electron attachment by : $\ln K = -\Delta G/RT$, or as a function of temperature, $d(\ln K)/d(1/T) = -\Delta H/R$. For electron attachment reactions the exothermicity of reaction is approximately equal to the electron affinity of the neutral molecule X. Therefore, by plotting the relative rate coefficients obtained as a function of

Figure 4. Electron density versus reaction distance when electrons are lost by diffusion only, lost by attachment to an attaching gas and by diffusion, and lost by attachment to an attaching gas and by diffusion but gained by detachment of negative ions.

1/T, a van't Hoff plot, we can determine the electron affinity. In principle the intercept of this plot should give the entropy of reaction. For our limited temperature range, this value will not be accurate. Analyzing our data for pentafluoropyridine and octafluorotoluene in this way, Figures 5 and 6, we obtained values of 0.83 eV and 0.72 eV, respectively. These values compare favorably with values estimated from the ion-molecule charge transfer measurements by Kebarle and coworkers (see Kebarle and Chowdhury[4]) who found values of 0.68 eV for pentafluoropyridine[4] and 0.83 eV for octafluorotoluene. [5] It is important to note that the relative electron affinity values for these two molecules obtained with our electron attachment-detachment data differ from those of Kebarle and coworkers by sign. We suggest that this difference arises from the lack of precision of our values; the experiment of Kebarle and coworkers directly monitors relative rates of electron transfer.

In order to determine whether this analysis is appropriate, we began exploring the implications of different microscopic mechanisms for electron detachment. As with associative attachment, such an interpretation as presented above, requires that the rate coefficients measured are pressure independent. Preliminary data on attachment and detachment to octafluorotoluene showed no pressure dependence. However, we found a pressure dependence for the rate coefficients for electron attachment to $c\text{-}C_4F_8$ and detachment from $c\text{-}C_4F_8$, a system previously investigated by Miller et al.[3] Interpretation of this pressure dependence was not straight forward. Extensive studies must be performed which quantitatively probe the effects of pressure, temperature, reactant neutral flow rate, electron temperature, and reaction time dependence.

Dissociative Attachment Reactions

Chloroacetonitrile and pentafluorobenzene react with thermal electrons at 300K dissociatively. Only one dissociative product channel is operative in the reaction of electrons with chloroacetonitrile while it appears that at least 2 product channels are operative in the reaction with pentafluorobenzene, i.e., two different product ions are formed in the reaction.

Chloroacetonitrile ($ClCH_2CN$) attaches electrons dissociatively to form Cl^- over the full temperature range studied (295-556K). While the thermochemistry of chloroacetonitrile is not well established,

Figure 5. (a) Electron attachment rate coefficient for reaction with pentafluoropyridine as a function of temperature (b) Electron detachment rate coefficient as a function of temperature (c) Equilibrium constant for electron attachment and detachment as defined in the text.

20-14

(a)

(b) C7F8 DETACHMENT RATE vs. TEMPERATURE

(c) C7F8 ln KEQ vs. I/TEMPERATURE

$\Delta H = 0.72 eV$

Figure 6. (a) Electron attachment rate coefficient for reaction with octafluorotoluene as a function of temperature (b) Electron detachment rate coefficient as a function of temperature (c) Equilibrium constant for electron attachment and detachment as defined in the text.

estimates[6] of the heat of formation of $ClCH_2CN$ indicate that formation of Cl^- + CH_2CN is the only

exothermic (and exoergic) dissociative reaction channel possible with $\Delta H_{rxn} \sim$ -68 kJ mol^{-1}. No

information is available on the electron affinity of chloroacetonitrile so that we cannot assign an enthalpy

for associative attachment.

The rate coefficient for electron attachment to chloroacetonitrile is moderate at 295K, 4 x 10^{-8}

cm^3 s^{-1}, and increases rapidly to 1 x 10^{-7} cm^3 s^{-1} at 556K. Since the reaction is exoergic, we presume that

the positive temperature dependence arises from an activation barrier to dissociative attachment. A

simple Arrhenius analysis of this temperature dependence, gives a pre-exponential of 2.4 x 10-7 cm^3 s^{-1}

and an activation energy for this exothermic process of ~5 kJ mol^{-1}. In this model, the pre-exponential

factor can be considered a collision rate coefficient of sorts. Interestingly, this experimental "collision rate

coefficient" is approximately the value estimated using the Klots approximation for the true collision rate

coefficient. Despite the appeal of this simple analysis, it is probably not appropriate; the Arrhenius plot of

lnk vs. 1/T is not linear over our experimental temperature range. The problem with this simple analysis is

that it assumes a single rate limiting and temperature dependent process is occurring. Based on current

understanding for electron attachment reactions, we presume that the parent negative ion is formed

initially, independent of what product channel is observed, although the state of the parent negative ion

cannot be assigned. In other words, the mechanism for dissociative electron attachment to

chloroacetonitrile forming Cl^- + CH_2CN can be described by:

$$e + ClCH_2CN + M \Rightarrow [ClCH_2CN^-]^* + M$$, where M is a collision partner (=He in our experiment)

$$\Downarrow$$

$$Cl^- + CH_2CN$$

The temperature dependence observed, therefore, combines the dependence of both the associative

electron attachment process and the subsequent unimolecular decomposition of the internally excited

negative ion.

Pentafluorobenzene reacts with electrons at 300K dissociatively. Major peaks in the mass

spectrum were at m/e = 98, 148, 167, and 184. Investigation of the peak height dependence on reactant

neutral flow rate indicates that the m/e 184 peak arises from secondary ion-molecule reactions of the primary products. The other three peaks arise principly from electron attachment. The anions with m/e = 167 and 148 correspond to attachment of an electron with loss of an H atom and a HF molecule respectively. Loss of HF upon addition of an electron is analogous to the reaction of benzene with O^- forming the benzyne radical anion.[7] We believe this is the first observation of the $C_6F_4^-$ anion. The anion formed from electron attachment followed by hydrogen loss has been observed in several other experiments but it's source was not pentafluorobenzene, as in our experiments. We do not currently have a hypothesis as to the structure of the anion observed with m/e 98. Based on our peak height versus flow rate analysis, the positive ion spectrum, and the structure of our reactant, we suggest that this anion is formed from an impurity in our pentafluorobenzene sample. Further work is needed to support the product identifications proposed and to quantify the electron attachment rate coefficient.

Other Reactions Studied

We investigated the reaction of electrons with hexafluoropropyleneoxide and ClF. However, uncertainties as to the purity of these species combined with the complicated mass spectra observed made interpretation of these data impossible. It is interesting to note that an anion was observed at the m/e corresponding to the parent ion of hexafluoropropyleneoxide which would indicate that the parent ion is bound, i.e., it has a positive electron affinity.

Development of an Experimental Procedure and Analytical Strategy for Highly Reactive and Low Vapor Pressure Compounds

As discussed above, it is often difficult to accurately monitor the flow rate of compounds into the flow tube. Both species which are highly reactive and those with low vapor pressures pose such problems. In the case of highly reactive species we have estimated rate coefficients by extrapolating the measured rate coefficients as a function of the length of time the reactant neutral vapor had been flowing in the inlet lines. Such a procedure not only takes experimental time and wastes resources; it is not a satisfactory solution. Following the lead of Smith et al.[2] we decided to probe the utility of using a

combination of second order kinetics and pseudo first order kinetics to solve for both the reactions rate coefficient and the neutral reactant flow rate.

The procedure involves measuring the electron loss with distance under conditions where the reactant neutral is in excess (pseudo first order kinetics) and under conditions were the reactant neutral and electron densities are similar (second order kinetics). These two experiments must use the same neutral flow rate, i.e., the electron density must be changed between experiments and the experiments must be performed on a time scale short, as compared with the drift in flow rate observed. Analysis of the data obtained should lead to accurate and unambiguous values for both the electron attachment rate coefficient and the reactant neutral flow rate.

Application of this procedure is complicated by the experimentally and theoretically limited ability to vary electron density over a wide range of values. Furthermore, analysis depends on the resolution of differences between the first and second order kinetics data and these differences are generally small. Finally, an accurate analysis critically depends on an accurate measure of the diffusion rate which can change significantly in the presence of a high mass attaching gas.

We have analyzed data for electron attachment to several of our reactive/low vapor pressure compounds but are not convinced of the uniqueness of our rate coefficient and neutral concentration values. For comparison, we obtained data under the prescribed conditions for well-behaved electron attaching reactants and are currently analyzing these data. We hope to better define experimental parameters which will lead to unique, accurate and precise determinations of the attachment rate coefficient and the neutral reactant flow.

Conclusions and Suggestions for Future Research

Halogenated aromatic molecules tend to form stable parent negative ions upon reaction with low energy electrons and many do so efficiently. Of the aromatic species studied, only pentafluorobenzene reacts with electrons dissociatively rather than associatively. It is likely that in this case, the associative product anion is not bound. Dissociative electron attachment was also observed in the reaction with chloroacetonitrile. Despite the exothermicity of this reaction, the efficiency for electron attachment

increases with increasing temperature suggesting that at least one aspect of the dissociative attachment process involves an activation barrier.

Weakly bound anions may undergo electron detachment. Electron detachment from the anion of octafluorotoluene (for T$\geq$ 510K) and pentafluoropyridine (T$\geq$ 465K) was observed. We have used the data obtained to quantify the relative rates of electron attachment and detachment. However, further work is needed to understand the fundamental meaning of these values.

Finally, development of an analytical method to facilitate accurate determinations for highly reactive species and low vapor pressure species has been initiated. Such a method will greatly increase the range of species which may be studied with this otherwise very versatile technique.

Future work will focus on completing comprehensive investigations of each reaction, including temperature and pressure dependence studies of the electron attachment rate coefficients. Extensive studies will probe the mechanism of electron detachment observed in our experiment. These studies will involve the systematic variation of pressure, temperature, reactant neutral flow rate, electron temperature, and reaction time dependence. To further our chemical insight into electron attachment reactions we will investigate electron attachment to molecules chemically similar to those investigated to date. These data will help to further reveal the influence of chemical structure on the electron attachment process. Finally, we will make both experimental design changes and procedural changes in order to facilitate the determination of electron detachment rate coefficients to highly reactive and low vapor pressure species.

References

[1] Smith, D. and Spanel, P., Adv. Atomic Molec. & Optical Phys., **1994**, 32, 307.

[2] Smith, D.; Spanel, P. Mark, T.D., Chem. Phys. Lett., **1993**, 213, 202.

[3] Miller, T. M.; Morris, R. A.; Miller, A. E. S.; Viggiano, A. A.; Paulson, J. F., Int. J. Mass Spectrom. Ion Proc., in press.

[4] Dillow, G. W.; Kebarle, P., J. Am. Chem. Soc., **1989**, 111, 5592.

[5] Kebarle, P.; Chowdhury, P., Chem. Rev., **1987**, 87, 513.

[6] Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G., J. Chem. Phys. Ref. Data, **1988**, 17, Supplement No. 1.

[7] Guo, Y.; Grabowski, J. J., J. Am. Chem. Soc., **1991**, 113, 1.

Ionospheric Tomography Using a Model Based Transformation Maximum Entropy Technique

Walter S. Kuklinski
Associate Professor
Department of Electrical Engineering

University of Massachusetts Lowell
1 University Avenue
Lowell, MA 01854

Ionospheric Tomography Using a Model Based Transformation Maximum Entropy Technique

Walter S. Kuklinski
Associate Professor
Department of Electrical Engineering
University of Massachusetts Lowell

Abstract

Ionospheric tomography uses total electron content (TEC) data obtained from a series of ground-based receivers and orbiting beacon satellites to determine the corresponding ionospheric electron density. The geometry of this technique does yield enough information to uniquely determine the ionospheric electron density unless additional information or assumptions are utilized. In this work a solution technique that can utilize both ionospheric model information and an entropy weighting for variations of potential solutions from a given model ionosphere has been developed. Results from both simulated and actual TEC data have yielded reconstructions that closely match the associated electron densities.

Ionospheric Tomography Using a Model-Based Transformation Maximum Entropy Technique

Walter S. Kuklinski

Introduction

The application of tomographic techniques to the problem of determining ionospheric electron densities from TEC data, obtained from a series of ground-based receivers and orbiting beacon satellies, has created the possibility of wide-scale, low-cost ionospheric monitoring. The geometrical constraints imposed by the ground-based receiver and orbiting beacon satellite technique provides TEC data that contains limited information about the variation of electron density with altitude. One problem that all present ionospheric tomography methods [1] attempt to solve, is the multifaceted trade-off between determining what additional assumed characteristics of the ionospheric electron density will yield a unique ionospheric reconstruction; the degree to which that reconstruction closely approximates the actual electron density; the degree to which the assumed characteristics of the ionospheric electron density modify aspects of the reconstructed electron density that in fact are uniquely determined by the available TEC data; and the computational complexity of the resulting reconstruction algorithm. Two specific aspects of this problem have been addressed in detail however, and form the basis of the ionospheric tomography algorithm presented here. Raymund et al. [2] present an ionospheric tomography algorithm that assures a solution that is consistent with the TEC data and preserves all aspects of the reconstructed electron density that are uniquely determined by the available TEC data. Fougere [3] has presented a maximum entropy based ionospheric tomography algorithm that yields unique reconstructions using a limited set of assumptions about the ionospheric electron density. The specifics of these techniques will be reviewed and a new class of model based maximum entropy ionospheric reconstruction algorithms will be presented. Initial results obtained using this method will be presented along with suggestions for futher research.

Algorithm Development

The relationship between TEC data and electron density is assumed to be linear and described as:

$$\underline{c} = A\underline{x}$$ (1)

Where $\underline{c}$ is an (n x1) vector of TEC data, $\underline{x}$ is an (k x 1) vector of electron density data, and A an (n x k) matrix with elements that are the lengths of the rays paths for the nth TEC measurement through the kth cell of the electron density array. In a typical case n < k , ie. there are fewer measurements that electron density values, yielding an under-determined system that may not have a unique solution. Even if more measurements were available, the geometry limits the number of independent measurements such that the rank of **A** is always less than m, precluding a direct unique solution for $\underline{c}$.

The solution method proposed by Raymund el al. **[2]** utilized the matrices P_N and $P_N^{\perp}$ which represent projections into the nullity of **A** and the orthocomplement of the nullity of **A** respectively. These matrices can be determined via a singular value decomposition of **A**, and can also be used to compute a pseudo-inverse $A^{\dagger}$ that satisfies:

$$P_N^{\perp} \underline{x} = A^{\dagger} \underline{c}$$ (2)

The pseudo-inverse can be used to obtain a set of unique solutions each of which exactly satisfies the TEC data as :

$$\underline{x} = A^{\dagger} \underline{c} + P_N \underline{x}$$ (3)

Since the contribution to the solution from the second term is from the nullspace of **A** and is unobservable from the measurements, it   provides the mechansism to generate unique solutions that exaclty satisfy the TEC data. To produce a specific solution additional constraints must be specified. Raymund et al. **[2]** utilized a set of model electron densities and required that; $\underline{x}$ satisfy (3); be of minimum norm; and be a linear combination of the model electron densities.

21 - 4

A constraint that has also been used in ionospheric tomography, Fougere [3] as well as other image restoration problems, Gull and Daniels [4], is maximum entropy. In this method the entropy of the reconstruction defined as :

$$H = -(\underline{x}/S)^t * Log(\underline{x}/S) \qquad (4)$$

is maximized. In this formulation Log ( ) represents an element by element operation yielding a (k x 1) vector, in a similar manner, division by the scalar S is an element by element operation also yielding an (k x 1) vector. The specific manner in which S is obtained has a significant effect or the objective function used in a subsequent optimization process. If the vector is normalized using the sum of its' elements, ie. assuming the vector is isomorphic with a probability density function, the resulting objective function does not have a unique maximum. The contour plot in the upper left hand panel of figure 1. shows contours of equal entropy for a two dimensional vector. As can be seen, any vector with equal elements has maximum entropy. Using a fixed constant S, in this example a model electron density $\underline{m}$=[45 5]^t produced an S value of 50, rather than the sum of the vector elements produces a normalized entropy function. Contours of equal normalized entropy are shown in the upper right panel of figure 1. The normalized entropy has a unique maximum at $\underline{x}_{maxent}$ = [S/e S/e . . . S/e]^t and could be used to obtain a unique reconstruction that is consistent with the measured TEC data by maximizing $H_{\underline{x}}$ with respect to $\underline{x}$, where $H_{\underline{x}}$ is :

$$H_{\underline{x}} = -(A^\dagger \underline{c}/S + P_N \underline{x}/S)^t *log(A^\dagger \underline{c}/S + P_N \underline{x}/S) \qquad (5)$$

This would produce a reconstructed electron density that was as close to $\underline{x}_{maxent}$ as the TEC data would allow.

While this solution would utilize the model electron density, it would do so in a minimal way, that is only using the sum of the assumed individual electron density values, and in a manner that would bias the solution towards $\underline{x}_{maxent}$ rather than a "mean" solution $\underline{x}_{mean}$ = [S/k S/k . . . S/k]^t. In the reconstruction algorithm developed here, the model electron density $\underline{m}$ is used to determine a transformation T{ $\underline{x}$ }, that maps (k x 1) solution vectors $\underline{x}$, so that T{ $\underline{m}$ } = $\underline{x}_{maxent}$. The reconstruction is obtained by maximizing H_T with respect to $\underline{x}$, where $H_T(\underline{x})$ is :

$$H_T(\underline{x}) = -(T\{A^\dagger \underline{c} /S + P_N \underline{x} /S\})^t \, {}^*\log(\ T\{A^\dagger \underline{c} /S + P_N \underline{x} /S\}\ ) \qquad (6)$$

The transformation T{ $\underline{x}$] is defined as :

$$T\{ \underline{x} \} = \underline{x} + \underline{x}^t \underline{x} \underline{d} \qquad (7)$$

where :

$$\underline{d} = (1/\underline{m}\ \underline{m}^t)\ [\ S/e - m_1\ \ S/e - m_2\ \ .\ \ .\ \ .\ \ S/e - m_k\ ]^t \qquad (8)$$

The contour plot in the lower left hand panel of figure 1. shows contours of equal entropy for transformed two dimensional vectors using the sum of the vector as S, and a model electron density $\underline{m} = [45\ 5]^t$. Using this value of S the transformation yields maximum entropy values for all vectors $\underline{x} = \alpha\underline{m}$. The contour plot in the lower right hand panel of figure 1 was produced using the same model electron density with S equal to 50, the sum of $\underline{m}$. This transformation maps vectors such that a unique maximum exists at $\underline{x} = \underline{m}$. This transformation also allows variations in potential solution vectors proportional to the amplitude of those components in the model electron density, and tends to favor smooth solutions as indicated by the orientation of the equal entropy contours toward the diagonal as seen in the normalized transformed entropy contour plot panel of figure 1.

One drawback of a reconstruction algorithm based on the direct optimization of $H_{T(\underline{x})}$ with respect to $\underline{x}$, is the possibility of the negative electron density values. The contours of high entropy in the region around $\underline{m}$, seen in the normalized transformed entropy contour plot panel of figure 1, extend into the adjacent negative electron density region. The entire region around the assumed model electron density is mapped towards the maximum entropy point with no penalty for negative electron densities. Since any reasonable model electron density will have small or zero values, for example at the bottom or top of the ionosphere, some negative penalty or correction needs to be employed. In the algorithm presented here an iteratitive correction technique is used.

The model electron density is considered an initial estimate of the solution since for cases where $\underline{x} = \underline{m}$, maximizing $H_T(\underline{x})$ with respect to $\underline{x}$ will yield exactly $\underline{m}$. The negative

21 - 6

correction, $\underline{x}_{i+1} = P(\underline{x}_i)$, is performed by setting any negative values of the initial reconstruction to zero and using the resulting positive electron density as the model electron density for a subsequent reconstruction. This process is outlined in follows:

1. Initialize model as : $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \underline{m}_0 = \underline{m}$

2. Determine initial solution $\underline{x}_0$ via: $\quad\quad\quad \nabla_{\underline{x}}\left[H_T(\underline{x}_0)\right] = \underline{0}$

3. Update solution via negative correction: $\quad \underline{x}_1 = P(\underline{x}_0)$

4. Done if: $\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \underline{x}_1 = \underline{x}_0$

5. If $\underline{x}_1 \neq \underline{x}_0$ update model as : $\quad\quad\quad\quad \underline{m}_1 = \underline{x}_1$

6. Return to 2. where in general:

Determine updated solution $\underline{x}_i$ via: $\quad\quad \nabla_{\underline{x}}\left[H_T(\underline{x}_i)\right] = \underline{0}$

In all cases this process converged within a few iterations to a consistent all positive electron density.

Results

The utility of the normalized transformed entropy reconstruction algorithm was determined using a set of simulated TEC data. These data sets where generated for a geometry with four recievers spaced at 10^0 latitude increments and a satellite passing directly overhead at an altitude of 1000 km. The ionosphere was represented using 30 equal divisions in latitude and, 8 divisions in altitude 100 km in thickness starting at 100 km altitude. The TEC data was obtained for 23 uniformly spaced satellite positions resulting in 92 measurements. The contour plots seen in figure 2. show, a simulated model electron density in the upper panel, along with a simulated electron density in the lower panel. The case shown here has two enhanced regions at altitude level 6, latitude locations 6 and 24 respectively. These enhancements are approximately twice the background electron density. To obtain these reconstructions a congujate-gradient method was used to solve the optimization represented in the solution of $\nabla_{\underline{x}}\left[H_T(\underline{x}_i)\right] = \underline{0}$.

Contour plots of the pseudo-inverse and complete reconstruction for this case are seen in figure 3. Figures 4. and 5. are one dimensional line plots of the electron density (dashed line) and pseudo-inverse (solid line). In these plots each set of electron densities at a fixed altitude, starting the top of the ionosphere, is plotted vs increasing latitude. The algorithm was also tested using test case data from Raymund et al. [5]. The data consisted of TEC measurements, a model ionosphere and a corresponding incoherent radar scan. Figure 6. shows contour plots of the pseudo-inverse and the electron density measured via the incoherent radar scan in the upper and lower panels respectively. While figure 7. shows contour plots of the complete reconstruction and the electron density measured via the incoherent radar scan in the upper and lower panels respectively.

Discussion

The initial results obtained using the model based transformation technique presented here indicated the potential utility of this method for iononspheric tomography. Using both simulated and actual TEC data, restorations that closely approximated the corresponding electron densities where obtained. This was true ever for cases where the actual electron density had enhancements that where not indicated in any way by the models. The method utilized relative TEC data using the method outlined in [2], which modifies both the TEC measurements and the **A** matrix of equation (1) so as to allow the computation of both the electron density and TEC offsets.

One significant improvement, that would be required before an operational tomographic reconstruction algorithm is derived from this method, is the inclusion of a method to allow for the possibility of noise in the TEC measurements . While the actual noise in the TEC measurements is quite small, temporal variations in the ionosphere during the satellite pass can produce inconsistencies in the TEC data that can be modeled as additive noise. The present reconstruction algorithm produces reconstructions that yield TEC data that exactly equal the observed TEC data. Maximizing $H_{T(\underline{x})}$ with respect to $\underline{x}$ subject to the condition that the corresponding TEC data, $\underline{c}_{\underline{x}}$, satisfied:

$$\underline{c}_{\underline{x}} \: - \underline{c} \: \leq \: \underline{\varepsilon} \qquad\qquad (9)$$

where $\underline{\varepsilon}$ is a vector of allowable TEC variations, would produce reconstructions that took this source of error into consideration.

References

1. Raymund, T. D., Franke, S. J., and Yeh, T.C., Ionospheric tomography: Its limitations and reconstruction methods, J. Atmos. Terr. Phys., 56(5), 637-657, 1994.

2. Raymund, T. D., Bresler, Y., Anderson, D. N., and Daniell, R. E. , Model-assisted Ionospheric tomography: a new algorithm, To appear in Radio Science 1994 .

3. Fougere, P. F., Iononspheric radio-tomography using maximum entropy, part 1: Theory and simulation studies, In preparation.

4. Gull, S. F. and Daniell, G. J., Image reconstruction from incomplete and noisy data, Nature, 272, 686-690, 1978.

5. Raymund, T. D. Ionospheric tomography algorithms, Accepted for publication JIST, 1994
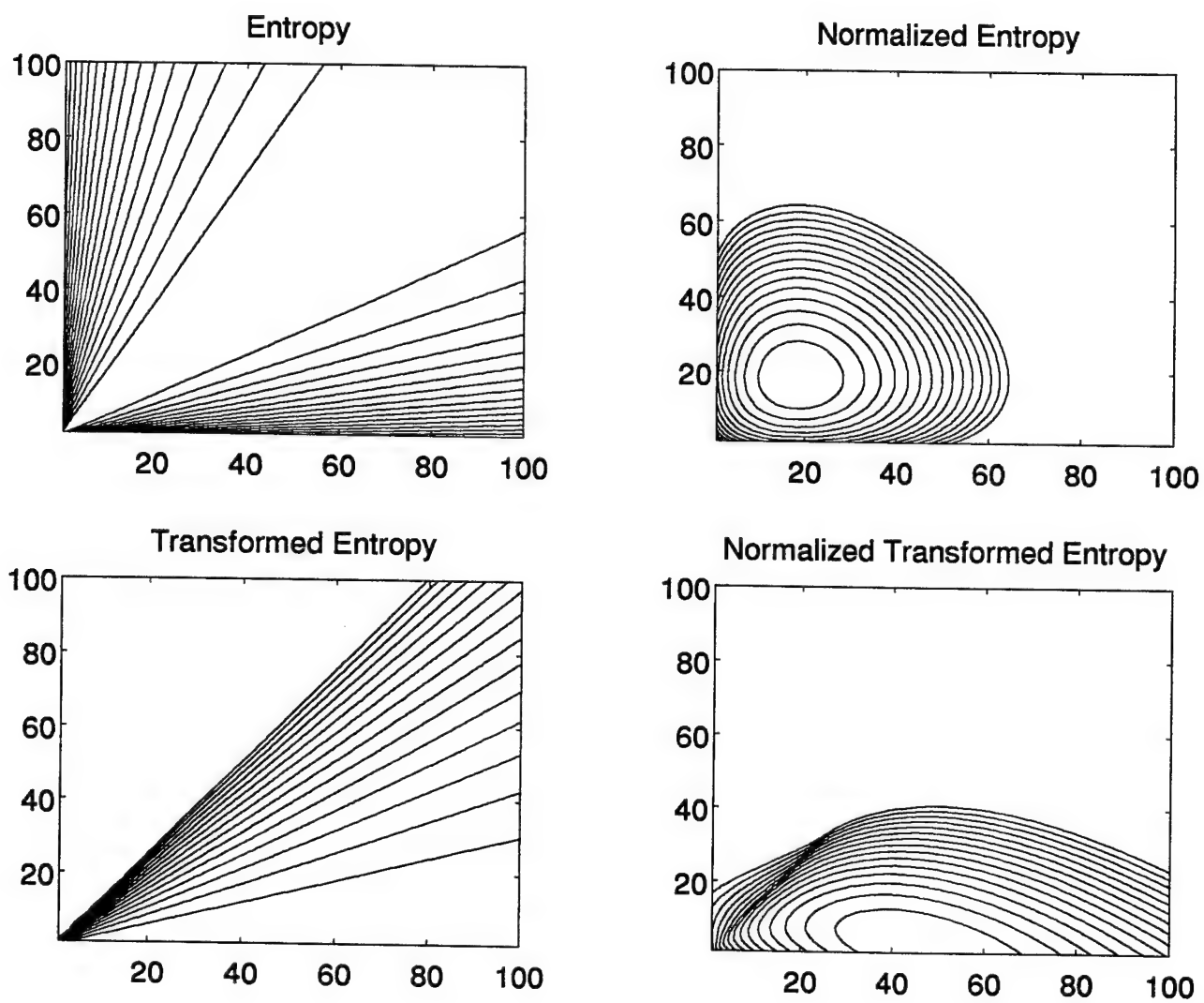
Figure 1

Figure 2

Figure 3

Pseudo Inverse vs. Electron Density

Figure 4

Reconstruction vs. Electron Density

Figure 5

Pseudo Inverse

Electron Density

Figure 6

Reconstruction via Model Method

Electron Density
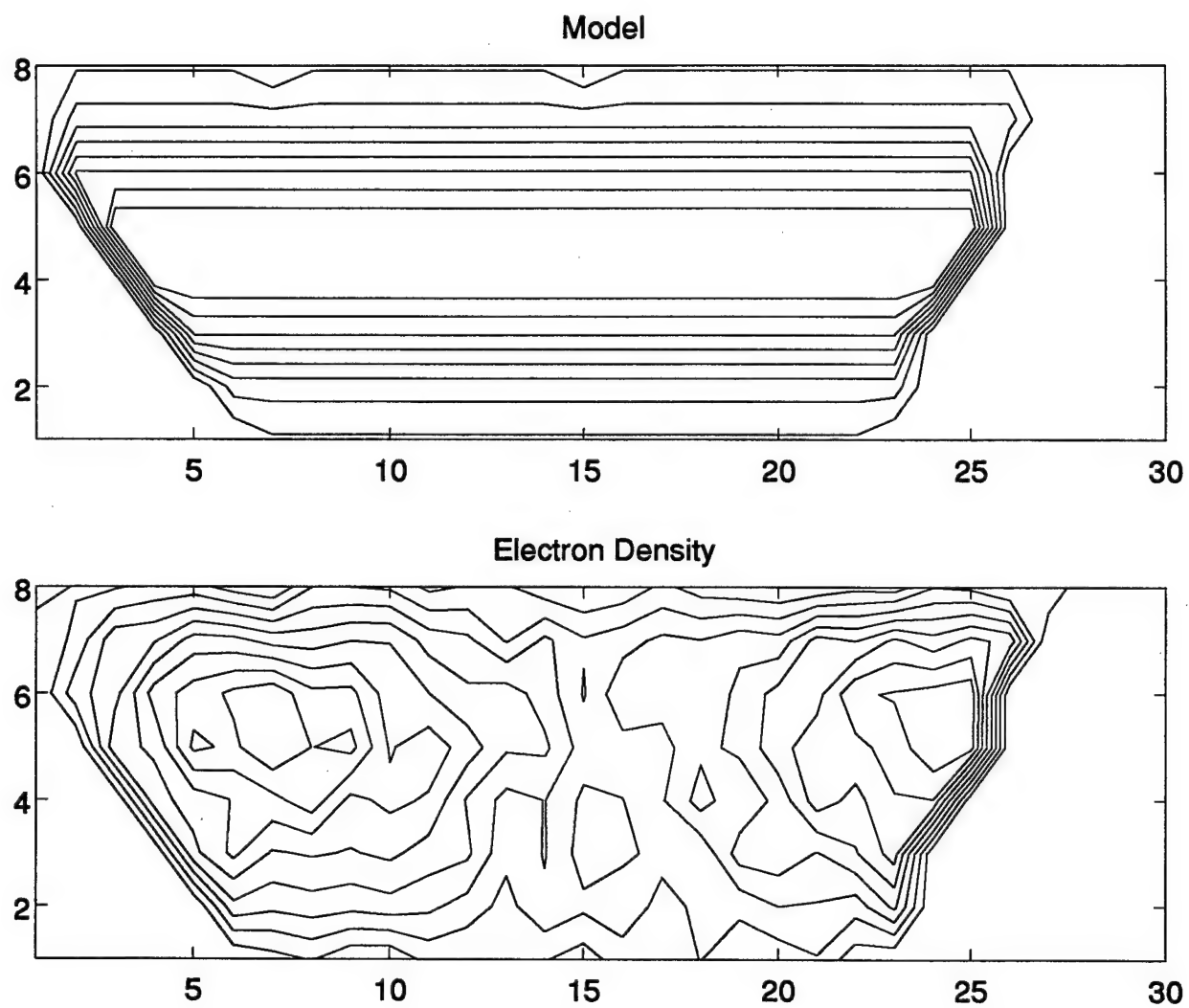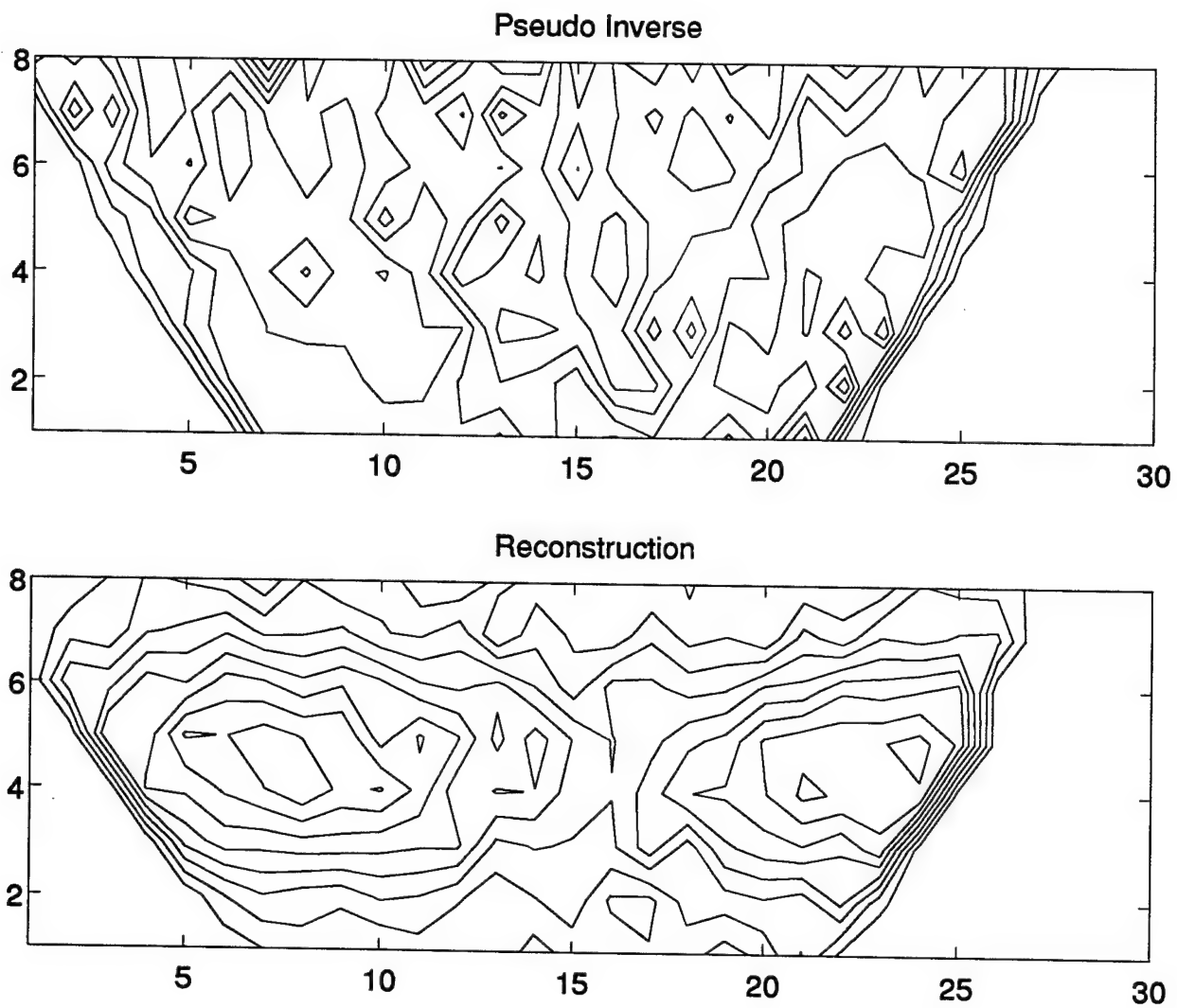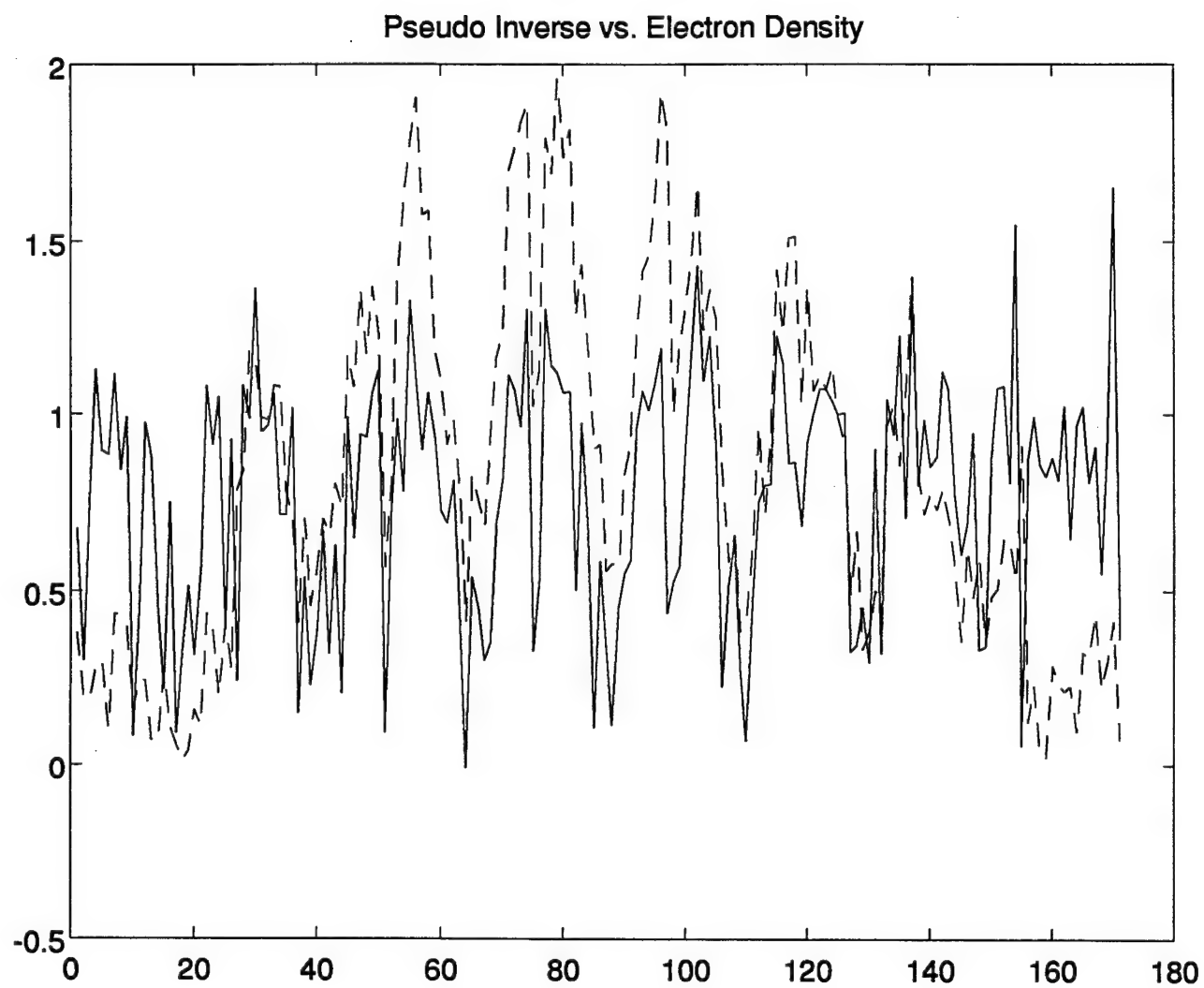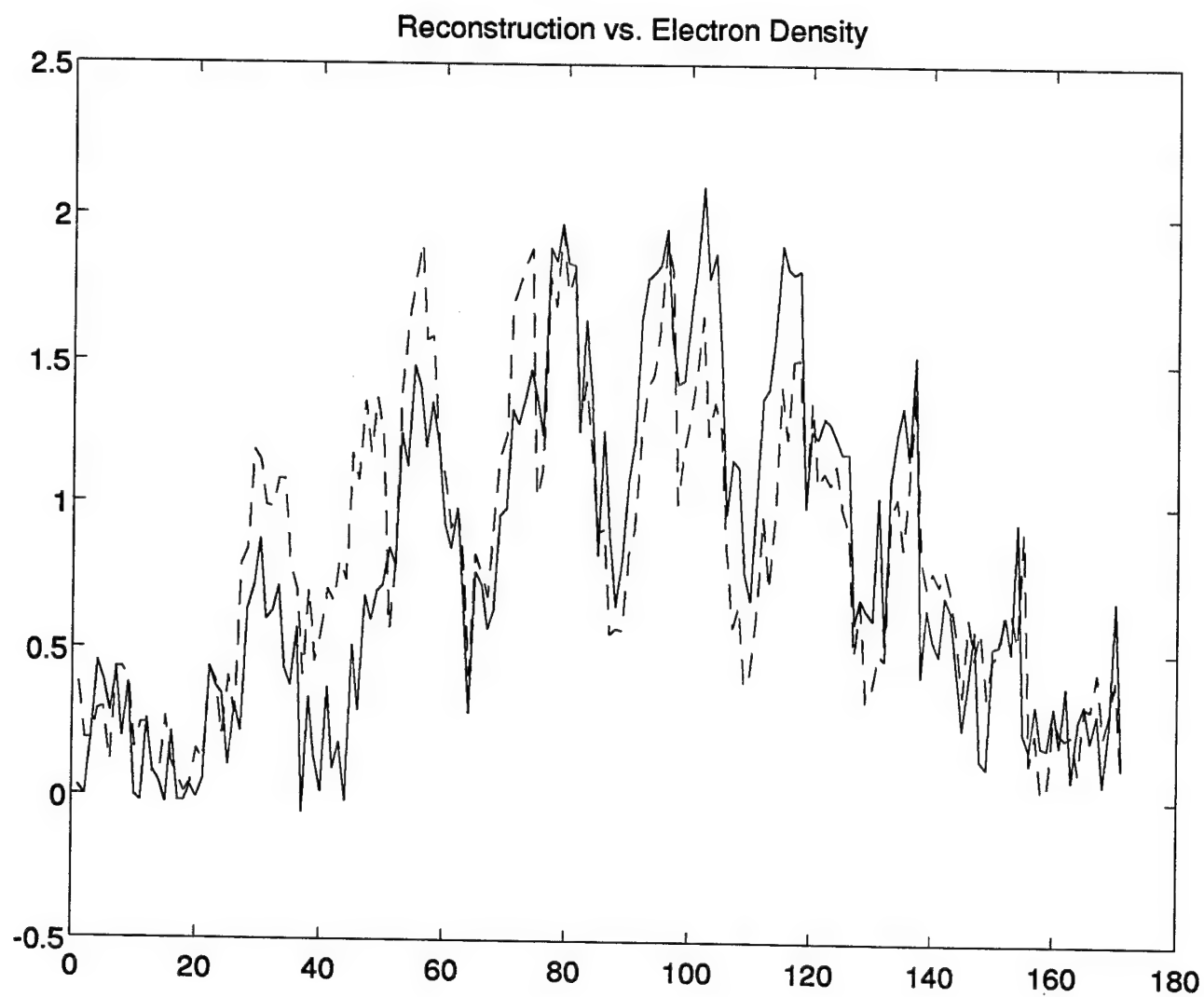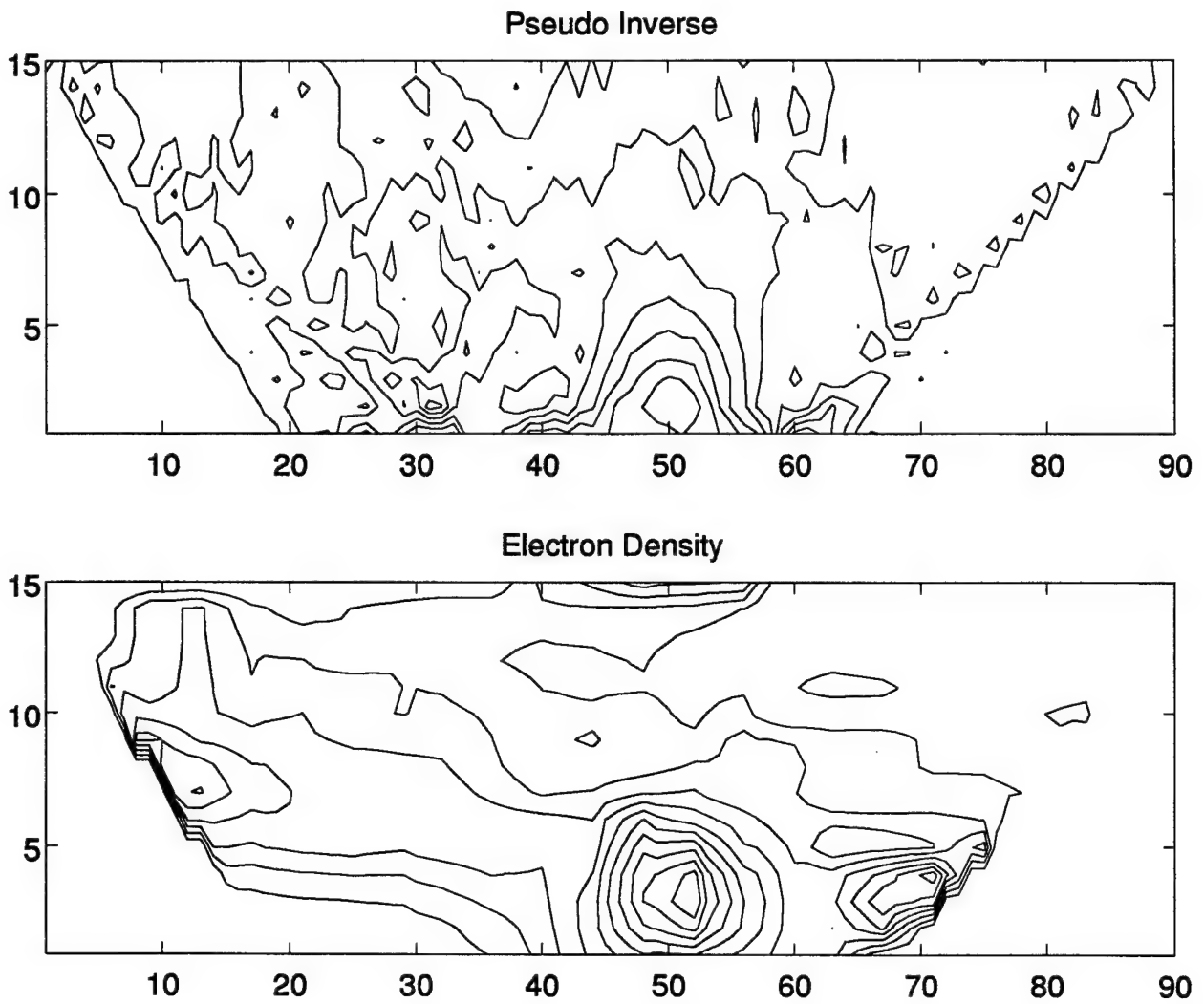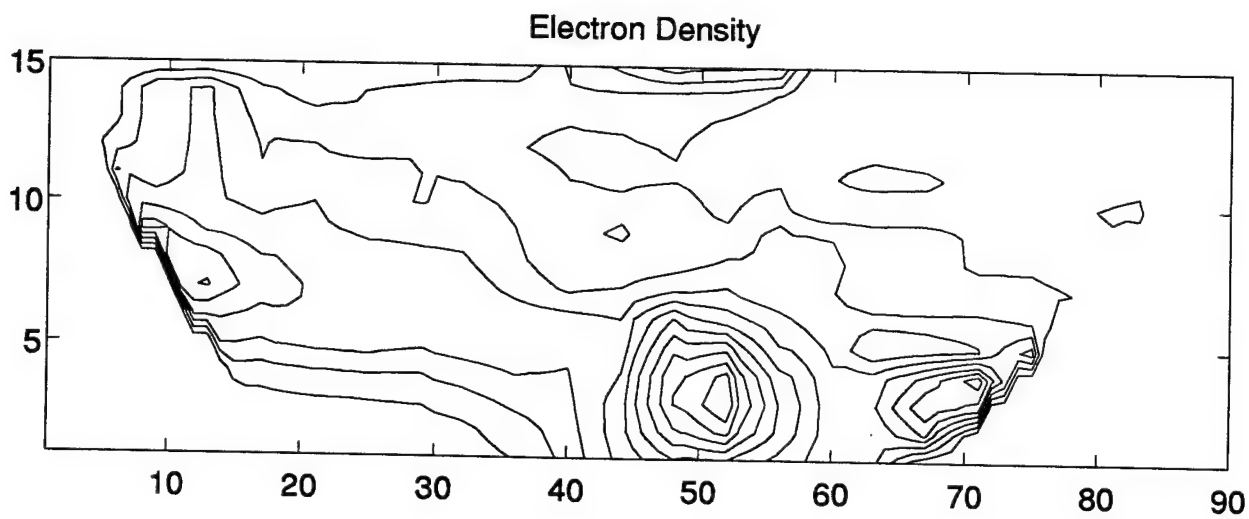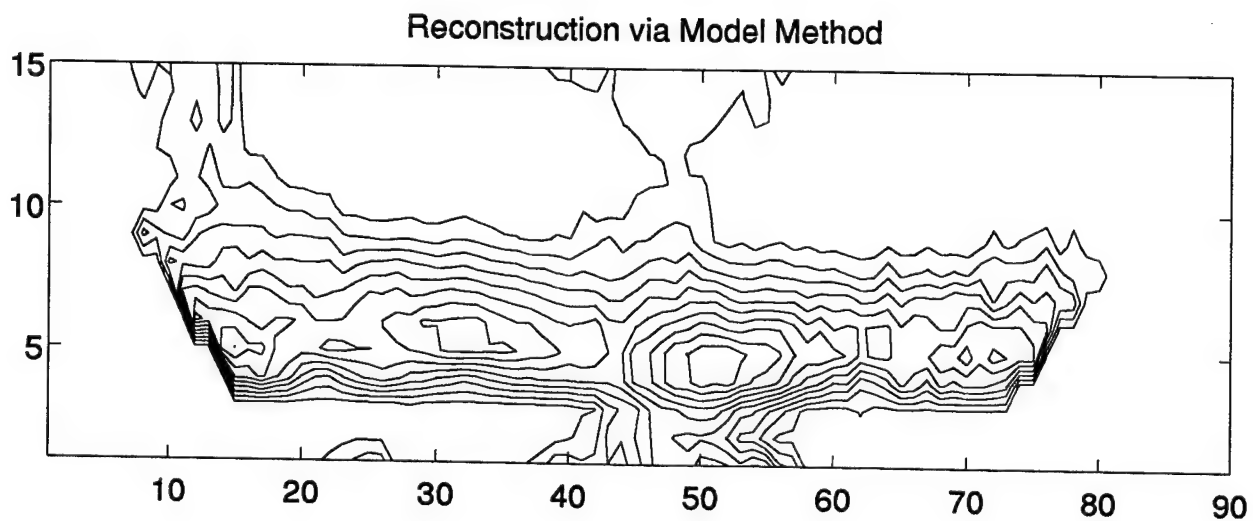
figure 7

Studies of Plasma Turbulence with Versatile Toroidal Facility
for Space Plasma Research

Min-Chang Lee

Plasma Fusion Center

Massachusetts Institute of Technology

Cambridge, Massachusetts 02139

Final Report for:

Summer Faculty Research Program

Phillips Laboratory

August 1994

Studies of Plasma Turbulence with Versatile Toroidal Facility
for Space Plasma Research

Min-Chang Lee
Plasma Fusion Center
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

We have conducted laboratory experiments with the Versatile Toroidal Facility (VTF) to investigate plasma turbulence, aimed at simulating the space plasma environment and cross-checking some space plasma experiments. Several students of mine including Dan Moriarty participated in the experiments. Dan Moriarty worked at the Geophysics Directorate of the Phillips Laboratory under the AFOSR sponsored Summer Graduate Student Research Program. VTF is a large plasma device which can generate magnetized plasmas with sharp density gradients and intense magnetic field-aligned currents. The VTF Plasmas have the key characteristics of the space plasmas, especially in the auroral region. Our experiments show that the VTF plasma turbulence is structured with both high-frequency and low-frequency wave modes. Such a plasma turbulence can be similarly produced by the sharp plasma density gradients and/or field-aligned currents in the ionospheric F region and in the topside ionosphere. The results of the VTF laboratory experiments are compared with those of the rocket experiments in space. We show that VTF can adequately simulate the naturally occurring plasma turbulence in the auroral ionosphere and complement the active plasma experiments in space.

Studies of Plasma Turbulence with the Versatile Toroidal Facility for Space Plasma Research

1. Introduction

Our research group at MIT Plasma Fusion Center has been actively conducting both theoretical and experimental studies of space plasma physics. In early 1989, my coworker, Prof. Ronald R. Parker and I [Lee and Parker, 1989, a research proposal submitted to AFOSR for funding] proposed that a large plasma chamber be constructed at the Nabisco Laboratory of the Plasma Fusion Center for research and education purposes [Lee, 1989, a proposal submitted to MIT Undergraduate Education Office]. The proposed laboratory experiments are aimed at simulating the space plasma environment and cross-checking results of our ionospheric plasma heating experiments at Arecibo, Puerto Rico. The construction of a toroidal plasma device, known as the Versatile Toroidal Facility (VTF) began in the summer of 1989 with a moderate budget.

Under the sponsorship of the Air Force Office of Scientific Research and the MIT Undergraduate Research Opportunities Program (UROP) office, the construction of VTF was made possible via two main sources. A majority of the construction materials such as the main toroidal field magnets, the poloidal field support structures, and the multi-megawatt power supplies were made available from discontinued fusion programs at the Oak Ridge National Laboratory (i.e., the ISX-B machine) and the MIT Plasma Fusion Center (i.e., the Tara Tandem Mirror mmachine). The second main resource was over 25 undergraduate and graduate students from various disciplines including physics, electrical, mechanical, chemical, and nuclear engineering.

The characteristics of VTF are illustrated in Section 2, showing how VTF can appropriately simulate space plasma environments, especially those of the auroral region. Described in Section 3 are our laboratory experiments conducted to investigate space plasma turbulence and its effects on radio wave propagation. The VTF experimental results are discussed and compared with those of rocket experiments in Section 4. Finally, presented in Section 5 are discussions and conclusions.

2. Characteristics of VTF

The Versatile Toroidal Facility (VTF) has a large toroidal vacuum chamber with a major radius, minor radius, and height of 0.9 meters, 0.35 meters, and 1.1 meters, respectively. It has 18 electromagnets mounted around the chamber generating toroidal fields up to 0.5 Tesla. There are 48 access ports - 16 each on the side, top and bottom of the chamber - for plasma diagnostics as illustrated in Figures 1(a) and 1(b).

Vertical magnetic fields up to 40 Gauss can be superimposed on the intense toroidal fields to form a helical magnetic field in the chamber [C. Yoo, M.S.Thesis, M.I.T., 1989]. The generation of helical magnetic fields serves two purposes: (1) to increase the plasma confinement time up to a mili-second, and (2) to guide electric currents flowing throughout the plasma chamber.

VTF can produce plasmas by either injected microwaves or electron beams. Microwaves at a frequency of 2.45 GHz are injected from a 3 kW CW magnetron. Hydrogen, the typical fill gas, is leaked into VTF to a working pressure around 7×10^{-5} Torr. Steady state plasmas produced by the microwaves via ECRH can have a peak density on the order of $10^{17} m^{-3}$.

The VTF electron beam system may provide seconds-long plasmas with densities as high as ten times that of ECRH plasmas. The principal elements of the electron beam system are two LaB_6 filaments installed at the bottom of the vacuum chamber. Each filament is heated to emit electrons that are subsequently accelerated by a 300 Volt potential applied between the filament and the vacuum chamber. Electrons flow upwards along the helical magnetic field producing field-aligned electric currents whose intensities can exceed 200 Amperes.

The geometry and characteristics of the VTF plasmas are delineated in Figures 2(a) and 2(b). Schematically shown in Figure 2(a) is the injection of microwaves and electron beams into the VTF plasma chamber. The density profile of the VTF plasmas produced by microwaves is illustrated in Figure 2(b). The peak density occurs at the location where the frequency of the incident microwave approximately matches the upper hybrid resonance frequency. Also illustrated in Figure 2(b) are magnetic field intensity and fractional density fluctuations versus the major radius of VTF. The helical magnetic field whose intensity is approximately equal to the toroidal field varies inversely with the major
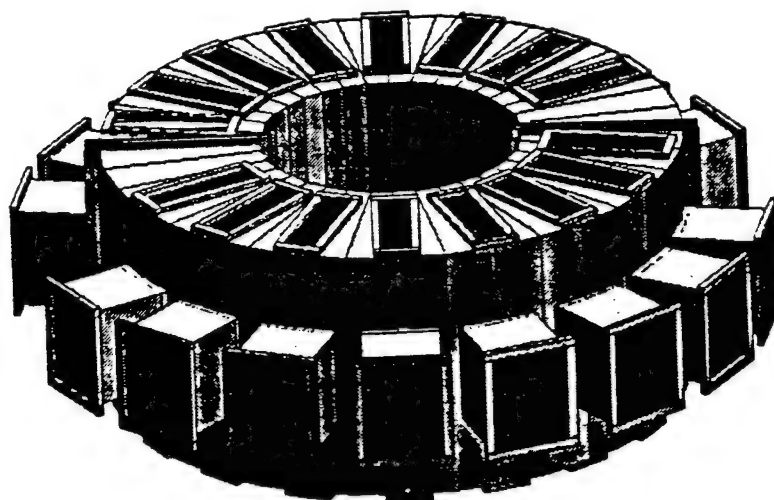
Figure 1(a). VTF vacuum chamber and side, top, and bottom ports.
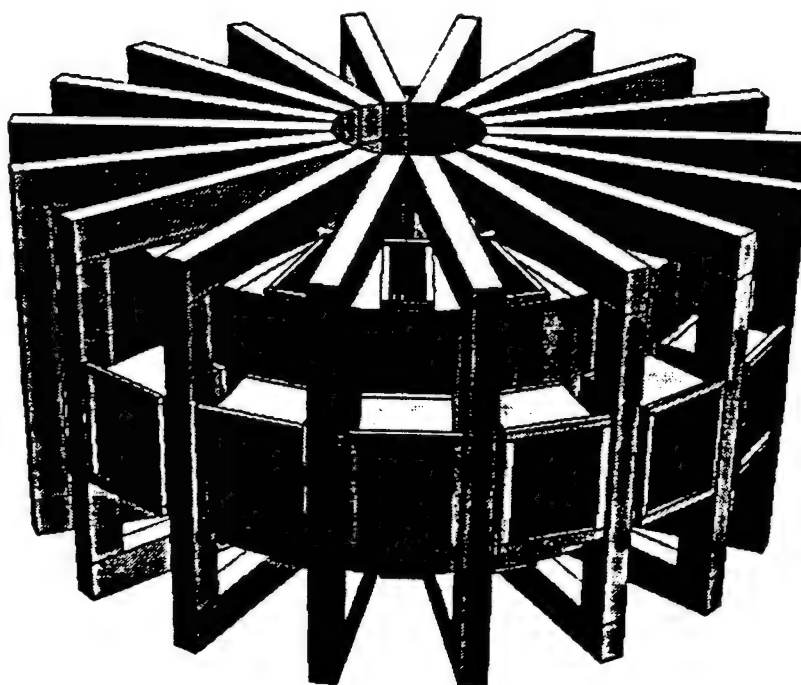
Figure 1(b). VTF vacuum chamber, TF coils, and (side, top, and bottom) ports.

22-5

(a)
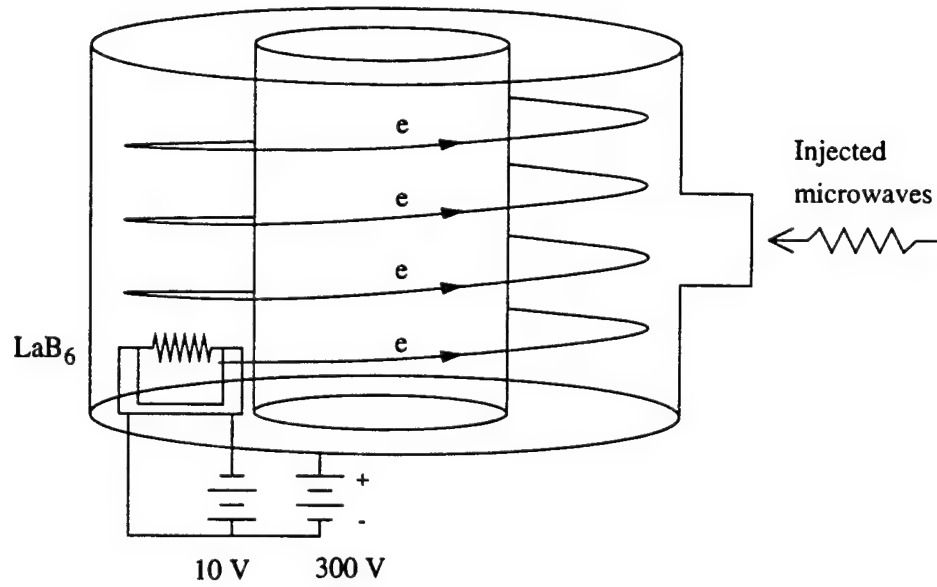
(b)
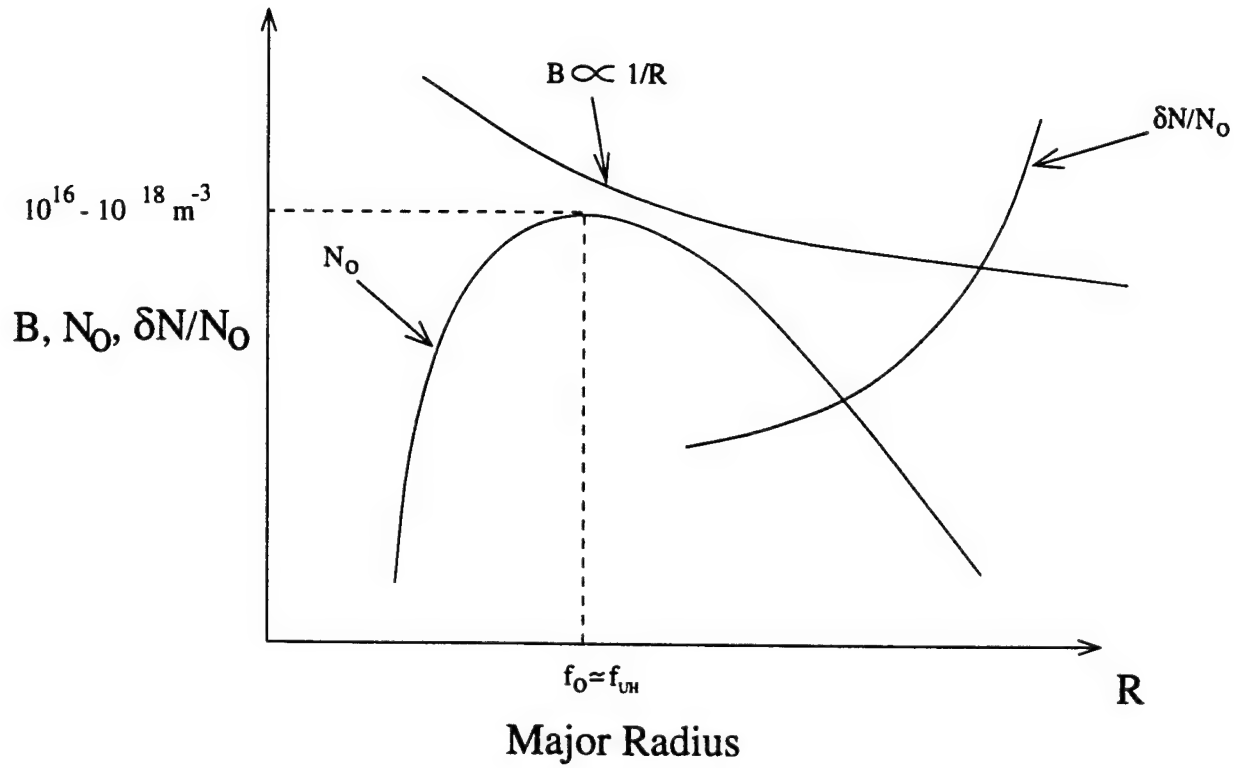
Figure 2(a) Injection of microwaves and electron beams into VTF;
2(b) Density profile, magnetic field, and fractional density
fluctuations in VTF.

radius. Large fractional density fluctuations ($> 10\%$) appear at the edges of the plasma density profile resulting from various plasma instabilities driven by field-aligned electric currents and sharp density gradients, as discussed in Section 4.

Several features of the VTF plasmas can reasonably simulate the space plasma environment. For instance, the turbulent auroral ionosphere is characterized by a sharp density gradient in the north-south direction and intense field-aligned electric currents. This auroral plasma condition can be simulated well by the VTF plasmas which have sharp density gradients in the radial direction and field-aligned electric currents in the azimuthal direction. The helical magnetic field of VTF is ideal for the studies of whistler waves that can interact with not only magnetospheric plasmas but also prominently with ionospheric plasmas [Lee and Kuo, 1984 (a) & (b); Groves et al., 1988; Liao et al., 1989; Dalkir et al., 1992].

3. VTF Experiments

Our laboratory experiments on plasma turbulence in VTF are aimed at investigating source mechanisms that cause the large fractional density fluctuations. The sharp density gradients and field-aligned electric currents are the potential sources of free energy generating plasma modes via instabilities. The excited plasma waves can be associated with intense density fluctuations. Because the VTF plasmas can simulate various space plasma environments, our investigation of VTF plasma turbulence is conducive to the understanding of space plasma turbulence.

The scenario of our experiments is as follows. With a hydrogen gas fill pressure of 7×10^{-5} Torr, a toroidal magnetic field of 876 Gauss, 3 kW of microwaves at 2.45 GHz injected in the O-mode create a plasma with a peak density near $7 \times 10^{16} m^{-3}$ at the location where the upper hybrid resonance frequency matches the microwave frequency. A typical microwave (RF) produced plasma density profile is shown in Figure 3(a). An electron beam generated by the heated LaB_6 filament at the bottom of the chamber was accelerated by a potential of -300 V and guided by the helical magnetic field to form field-aligned currents flowing throughout the chamber. The electron beam can cause additional ionization, enhancing the plasma density by nearly a factor of ten. Figures 3(b) and 3(c) show beam plasma densities for low and high currents, respectively.
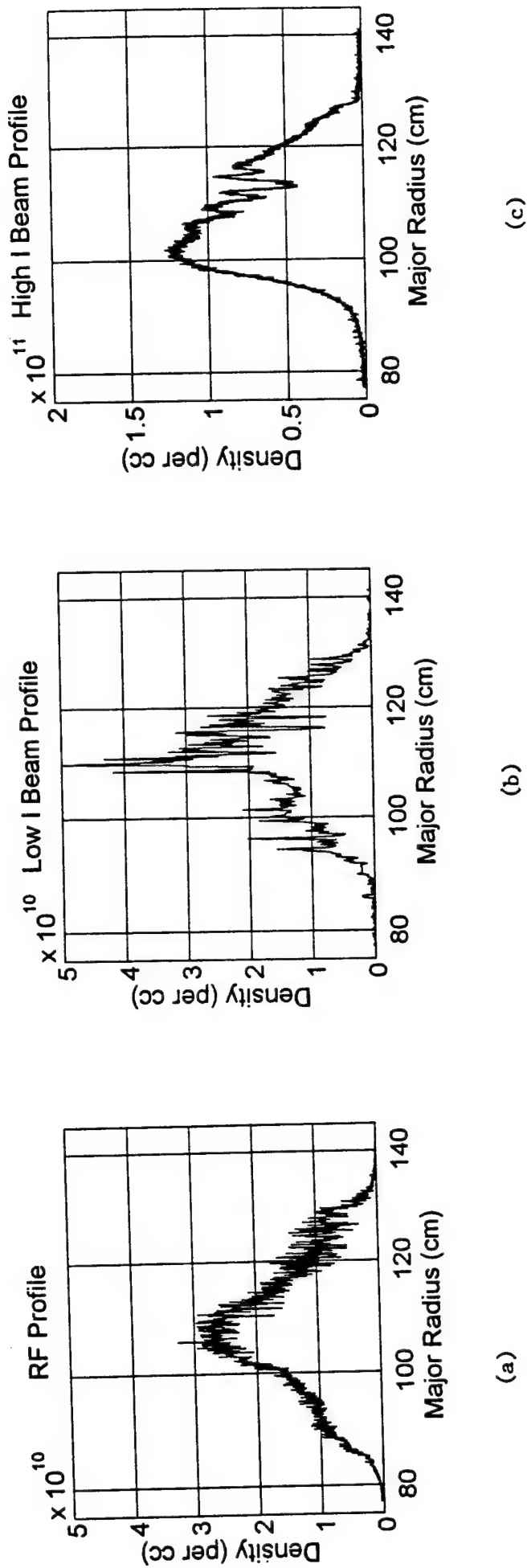
Figure 3(a) Microwave-produced plasma density profile in VTF;
3(b) Electron beam produced plasma density profile with lowe electric current;
3(c) Electron beam produced plasma density profile with high electric current.

Langmuir probes together with a spectrum analyzer were used to measure the excited plasma modes. Figure 4 provides a wideband view of all measurable modes in VTF with both the electron beam and microwaves as the plasma sources. We should point out that the excited modes have a broad range of frequencies up to about half of the electron gyrofrequency. Displayed in Figure 5 are the low frequency spectra of plasma modes corresponding to 3 cases wherein different electric currents (I=0, 15, 175 Amperes) were carried by the electron beam in addition to the injected microwaves. The ion plasma frequency (f_{pi}), the lower-hybrid resonance frequency (f_{lh}) and the ion gyrofrequency (f_{ci}) are marked in the figure, typically having values of 57 MHz, 40 MHz and 1.3 MHz respectively.

There are good reasons for us to speculate that several mechanisms play dominant roles in generating the spectra of the plasma modes in different frequency regimes. These mechanisms may be driven by field-aligned currents and/or sharp density gradients. For convenience, we divide the frequency spectra shown in Figure 4 into three regions labelled as (I), (II), and (III). Our spectral measurements indicate that "low-frequency" modes in region (I) can be preferentially excited at locations where sharp density gradients exist. By contrast, the excitation of the "medium-frequency" modes in region (II) and the "high-frequency" modes in region (III) is not sensitive to the density gradients. Spectral measurements of excited modes at different locations across the plasma were made with a radially scanning electrostatic probe. These results corroborate our speculation that different mechanisms contribute additively to the excitation of plasma modes in different frequency regimes.

In order to determine the dimensions of the field-aligned currents, measurements of plasma density profile across the magnetic field were made in both the radial and vertical directions. Plasma density along the magnetic field is uniform. The dimensions of the currents were estimated to be 0.25 meters (radial direction) by 0.05 meters (vertical direction) at the location of a Langmuir probe. This information will be used in the next section to examine source mechanisms that generate the turbulent VTF plasmas.

4. Comparison with Space Experiments

Experiments on the injection of electron beams into the ionosphere from sounding
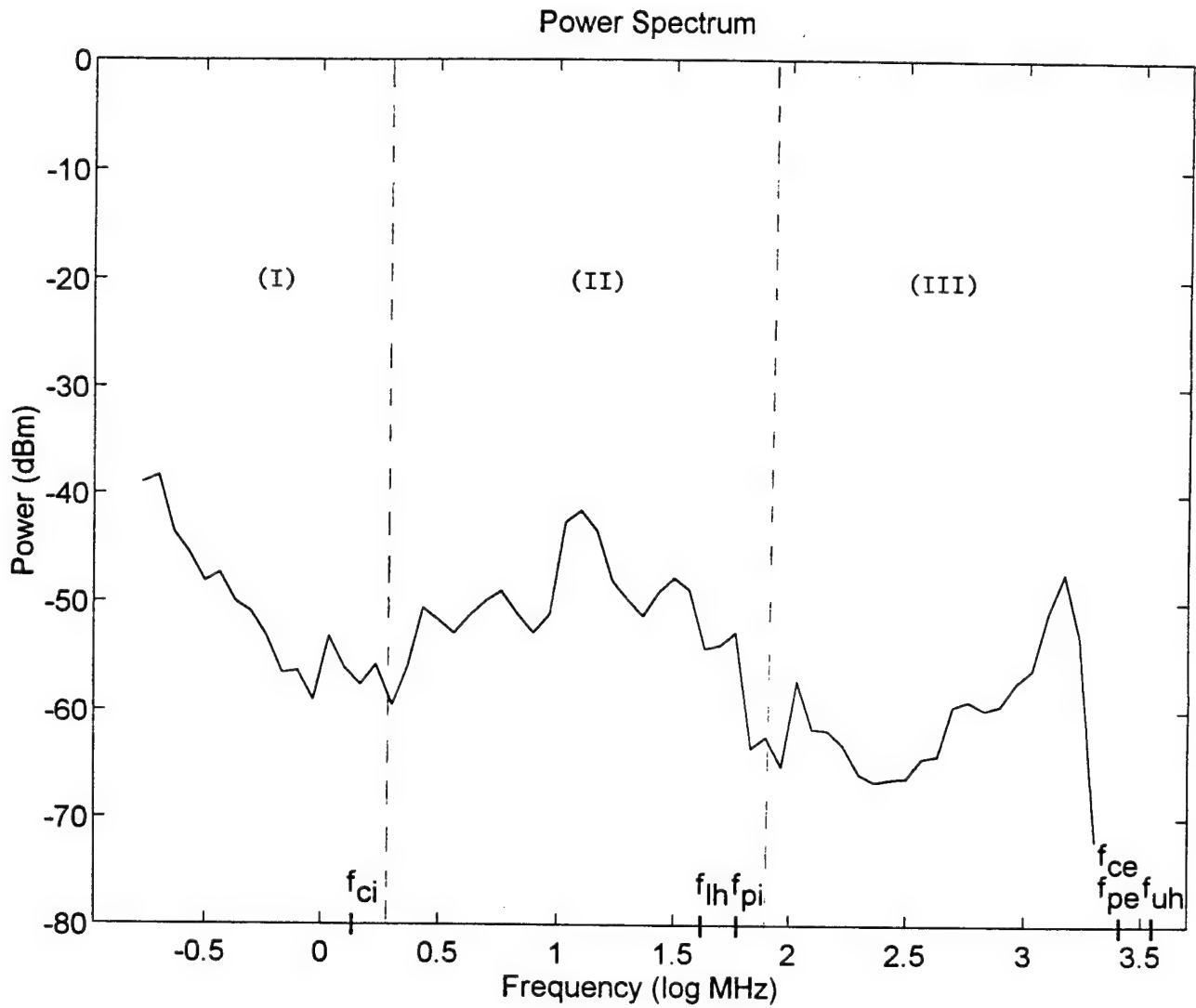
Figure 4. Spectra of plasma modes measured in the turbulent VTF plasmas
produced by both the injected microwaves and electron beams.
Marked with (I), (II), and (III) are regions representing
low-frequency, medium-frequency, and high-frequency regimes.
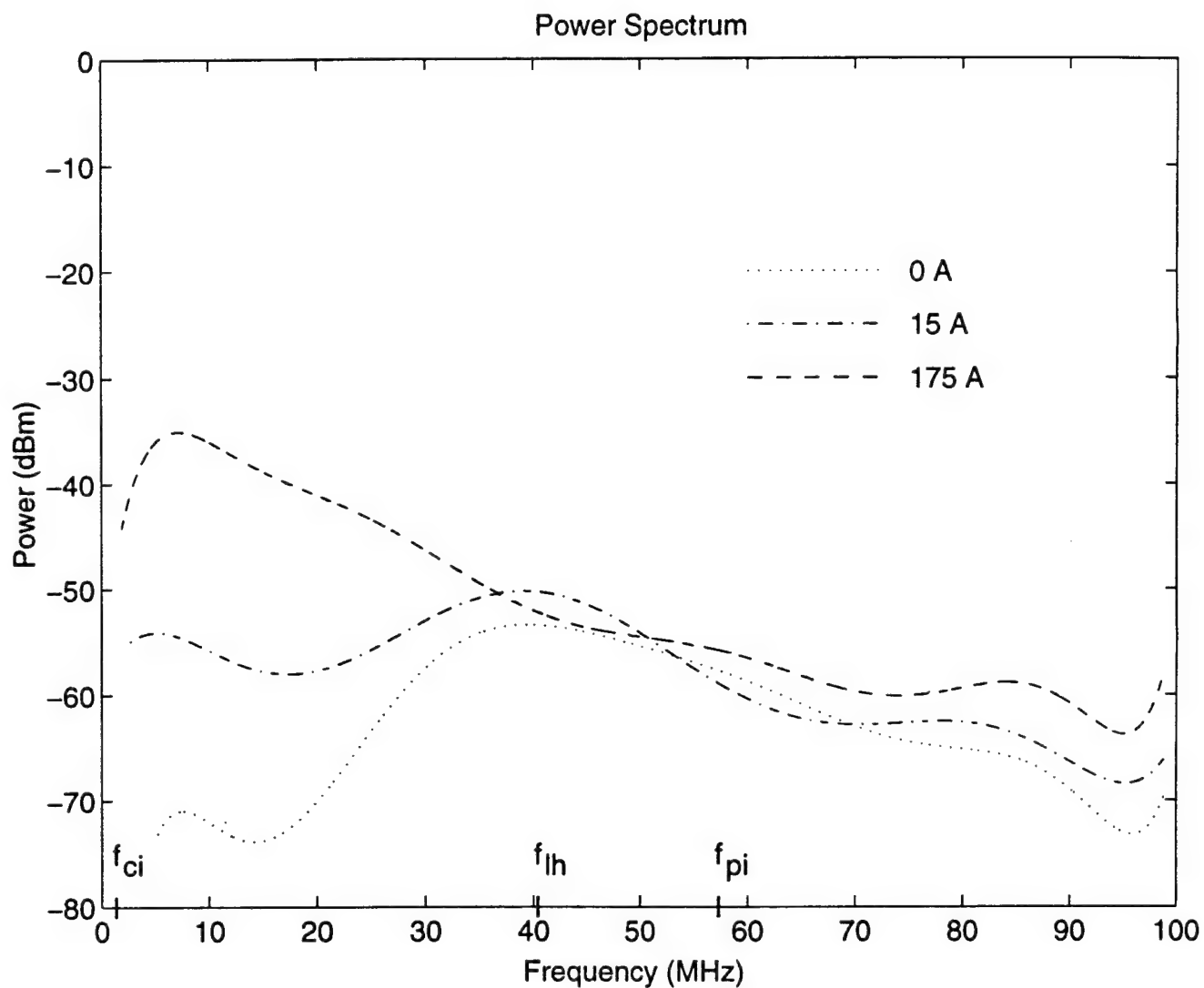
Figure 5. Illustration of excitation of lower frequency modes (primarily region (II), see Figure 4) by electric currents with different intensities.

rockets or space shuttles have been conducted by several groups to investigate beam-plasma interactions and wave-particle interactions in the auroral region [e.g., Winckler et. al., 1984; Kellogg et al., 1986; Gurnett et al., 1986; Reeves et al., 1988; Winckler et al., 1989; Winglee and Kellogg, 1990; Ginet and Ernstmeyer, 1991]. These experiments have shed light on the production of ionospheric plasma turbulencd by auroral electron beams. The geometry of the rocket experiments we will discuss here [Winckler et al., 1984; Winckler et al. 1989; Ginet and Ernstmeyer, 1991] is shown in Figure 6.

A plasma diagnostics package abbreviated PDP for short in the figure has two orthogonal sets of electric probes in the spin plane of the payload. A fifth probe projects upward along the PDP spin axis. These probes consist of insulated rods terminating in spherical sensors coated with graphite. The main payload (denoted by MAIN) injected two independent electron beams: one (gun 2) at a constant pitch angle of 100° or 110° and the other (gun 1) with a pitch angle nearly parallel to the magnetic field. Gun 1 operated at 36 kV and 250 mA, and gun2 was swept from 40 kV to 8 kV with a corresponding current in each 1 ms interval.

Presented in Figures 7(a) and 7(b) are typical electric field power spectra measured on the PDP during beam injection at a pitch angle of 110° [Winckler et al., 1989; Ginet and Ernstmeyer, 1991]. The plotted spectra include four different measurements of E_x and E_y denoted by a dashed line, solid line, dot-dashed line, and dotted line. Here, E_x and E_y refer to the electric field components measured by the two orthogonal sets of electric dipole antennas (see Figure 6). The dashed line represents the Fourier transform of the DC E_y component measured over a 200 ms interval every 0.4 ms. The solid line is the Fourier transform of the VLF broadband receiver measurement of E_y sampled every 0.05 ms over 100 ms. The dot-dashed line is the E_x component measured by the high swept frequency analyzer (HSFA) with a frequency resolution of 30 kHz and a sweep time of 102 ms between 60 kHz and 5.5 MHz. A composite background spectrum constructed in the same manner from data taken when there was no beam injection is shown by a dotted line.

Marked in Figure 7(b) on the top of the frame are the ion gyrofrequency (f_{ci}), the lower hybrid resonance frequency (f_{lh}), the ion plasma frequency (f_{pi}), the electron
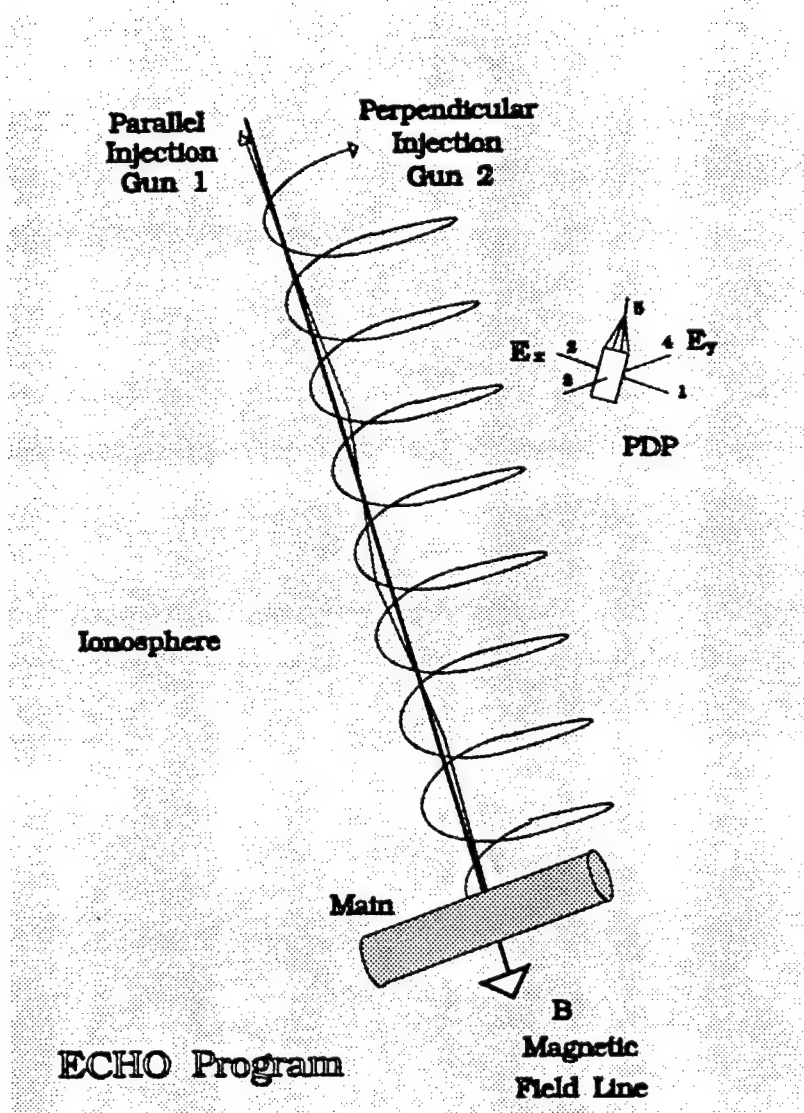
Figure 6. Geometry of electron beam injection experiments in space
with a rocket which has five electric probes on a plasma
diagnostics package (PDP) and two electron guns on the main
payload (MAIN) (Winckler et al., 1984).
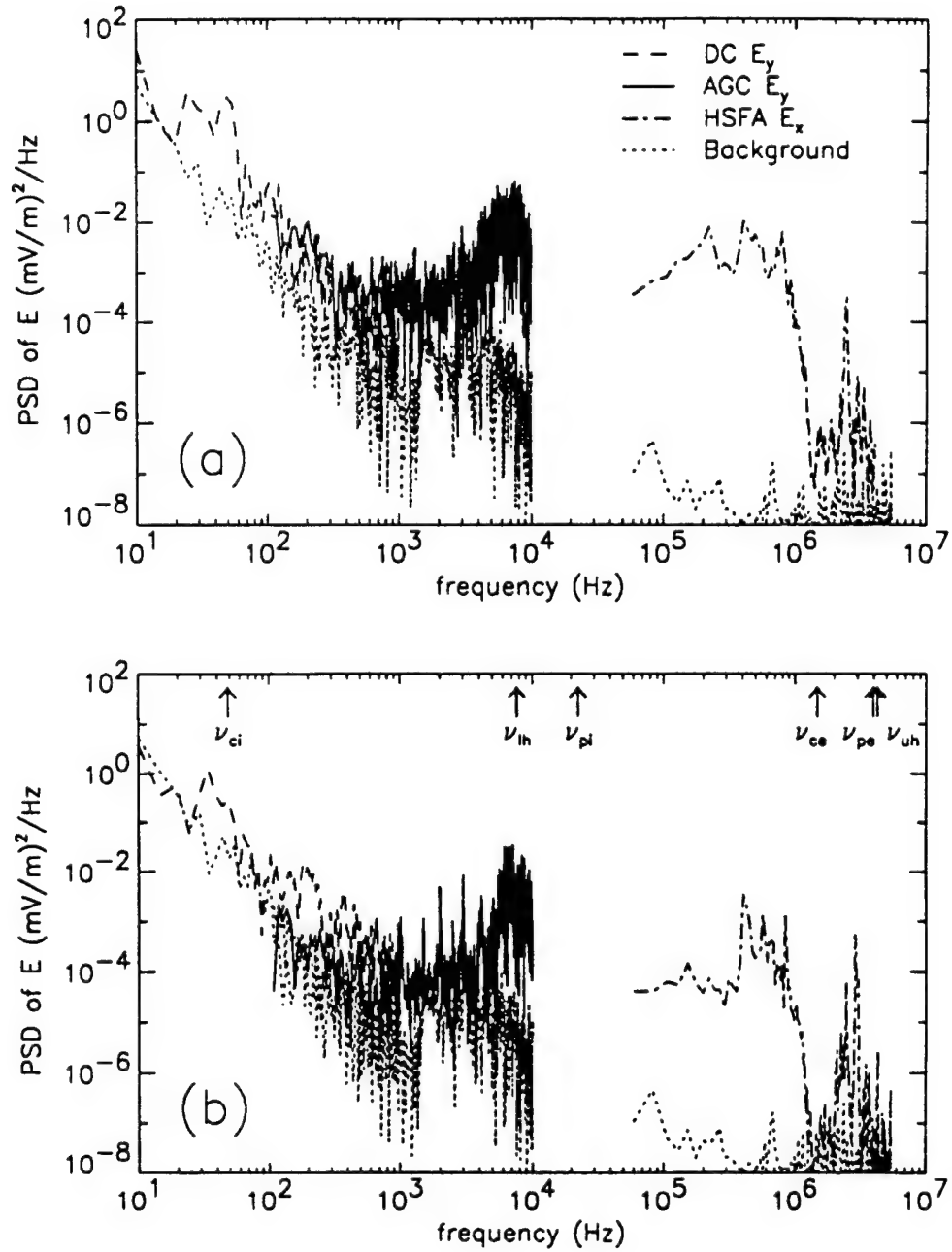
Figure 7. Typical electric field power spectra measured on the PDP associated with a 250 mA beam pulse injected at a pitch angle of 110° where (a) the beam energy is constant at 36 keV, and (b) the beam energy is swept through the range 40-8 keV every 1 ms (Ginet and Ernstmeyer, 1991).

gyrofrequency (f_{ce}), the electron plasma frequency (f_{pe}), and the upper hybrid frequency (f_{uh}). Compare the date from space experiments with those recorded in our laboratory experiments as displayed in Figure 4. It appears that they are quite similar. But careful examination of them as elucidated below indicates that different mechanisms operate in the space experiments and in the laboratory experiments.

First, high-frequency modes with frequencies exceeding the electron gyrofrequency (f_{ce}) were not excited in our VTF plasmas, while they were present in the beam injection experiments in space. This difference arises from the fact that 36 keV electron beams were injected in space experiments, whereas 300 eV electron beams were produced in the VTF plasmas. Note that the thermal electron energy in the ionosphere and in the VTF are 0.1 eV and 5-8 eV, respectively. In fact, the 300 eV electron beams reduced their speeds to nearly the electron thermal speed at the location of the probe via the ionization of neutrals along the beam path to the top of the chamber. Thus, the low-energy electron beams are unable to excite Langmuir waves and upper hybrid waves in VTF. By contrast, whistler waves (viz. modes in the frequency range: $f_{lh} < f < f_{ce}$) were produced in both space experiments and laboratory experiments possibly by the convective beam amplification of incoherent Cerenkov noise [Maggs, 1976].

One distinctive difference between the two experimental results is noted regarding the modes with frequencies near the lower hybrid resonance frequency (f_{lh}). These modes were favorably excited in the ionospheric plasmas by energetic electron beams, presumably by the modified two stream instability [Ginst and Ernstmeyer, 1991]. But the electron beams injected into VTF were not intense enough, as explained before, to cause the modified two stream instability. Although the modes at f_{lh} were not so highly peaked in the VTF plasmas, excited modes with frequencies adjacent to f_{lh} are clearly seen in Figure 5.

The other distinctive difference is noted regarding the low-frequency modes whose frequencies are less than the ion gyrofrequency (f_{ci}). One can see from Figures 7(a) and 7(b) that the ionospheric plasmas had large thermal noises in the low-frequency band (region I) and the medium-frequency band (region II). The energetic electron beams injected into the ionosphere cannot excite the low-frequency modes significantly, whereas

prominent excitation of these modes occurs in the VTF plasmas.

The characteristics of the excited waves by injected electron beams in the ionosphere and in VTF can be summarized as follows. Langmuir waves, upper hybrid waves, whistlers, and lower hybrid waves can be excited by energetic electron beams in the ionospheric plasmas. In the VTF plasmas, the injected electron beams produce whistler waves and especially low-frequency modes with frequencies less than the lower hybrid resonance frequency (f_{lh}) and even the ion gyrofrequency (f_{ci}).

5. Discussions and Conclusions

We pointed out earlier that the VTF was constructed to simulate the ionospheric plasma environment. Then how can we explain the discrepancies between the results of space experiments and those of VTF laboratory experiments? The excitation of high-frequency modes (i.e. Langmuir waves and upper hybrid waves) depends on the energy of the electron beams, as discussed before. This offers a reasonable explanation for the discrepancy only in the high-frequency regime of the wave spectra.

We believe that the discrepancy in the low-frequency regime of the spectra stems from the different background plasma environments. This fact can be understood from Figure 6. In the space experiments, the sounding rocket carried the electron guns as its main payload and the electric field probes on a plasma diagnostic package (PDP). During the experiments, the PDP separated at 1.5 m/s relative to the main payload. Thus the electron beam-plasma interactions and their subsequent diagnoses were carried out in a rather small ionospheric region which can be reasonably considered to be a uniform plasma.

By contrast, the VTF plasmas created by microwaves and/or electron beams have sharp density gradients across the magnetic field. They also have hot electrons but relatively cold ions. Hence, the VTF plasmas are inhomogeneous and have a large ratio of electron temperature (T_e) to ion temperature (T_i). Sharp density gradients and large T_e/T_i together with field-aligned electric currents are the important sources of free energy to excite ion acoustic waves and very low-frequency modes as explained below.

Kindel and Kennel (1971) suggest that field-aligned currents in the topside ionosphere

can excite plasma turbulence that may have the structure of electrostatic ion cyclotron waves (i.e., obliquely propagating ion Bernstein waves) and ion acoustic waves. Both kinds of waves can be favorably excited by electric currents with drift velocities smaller than the electron thermal velocity in plasmas with $T_e/T_i \geq 1$. The damping increment of the ion cyclotron waves is quite small even if the ion temperature is comparable to the electron temperature. This contrasts the ion acoustic waves which are highly damped when $T_i \simeq T_e$. Kindel and Kennel [1971] showed that when $T_e/T_i \gg 1$, say, 10, the threshold drift speed of the ion acoustic instability would be much less than that of the ion cyclotron instability. In the VTF plasmas, the electron drift velocity ($\sim 5 \times 10^4$ m/s) is less than the electron thermal velocity ($\sim 4 \times 10^6$ m/s) and the ion acoustic velocity ($\sim 4 \times 10^5$ m/s). However, the (T_e/T_i) of the VTF plasmas is estimated to be greater than 20. Therefore, the ion acoustic waves can be favorably excited in the VTF plasmas. It is, however, difficult to excite ion acoustic waves in the ionosphere. Although a large T_e/T_i was produced in the beam injection experiments in space [W.J. Burke, personal communication, 1994], the electron beam travelling in the heated ionospheric plasmas is different from an electric current flowing in the VTF plasmas. In the space experiments, energetic electron beam-plasma interactions occur to excite high-frequency wave modes as seen in Figures 7(a) and 7(b), while a low-energy electric current flows in the VTF plasmas to produce low-frequency modes.

The current convective instability was proposed by Ossakow and Chaturvedi [1979] as a potential mechanism to produce plasma turbulence with low-frequency modes in the auroral ionosphere. The auroral ionospheric plasma has a large density gradient in the horizontal plane (i.e., the North-South direction) and field-aligned currents flowing in the vertical direction. This plasma environment can be reasonably approximated by the VTF plasmas which have sharp density gradients in the radial direction and field-aligned currents in the azimuthal direction. Using the following VTF parameters: $T_e = 8$ eV, $T_i = 0.1$ eV, $f_{ce} = 2.45$ GHz, $f_{ci} = 1.3$ MHz, I = 10 Amperes, L = $n(\delta n)^{-1} = 0.1$ m, A (beam c/s) = 0.25 m x 0.05 m, and $k_{parallel}/k_{perpendicular} = 10^{-3}$, we find from Ossakow and Chaturvedi [1979] that the growth time of the convective instability in VTF is 0.1 msec. It indicates fast excitation of the current convective instability in the VTF

plasmas in comparison with the 1 ms confinement time. The frequencies of the excited modes are kV_d, where the electron drift velocity (V_d) is 5 x 10^4 m/s, which is less than the ion acoustic wave velocity, $V_s \sim 4x10^5$ m/s. We expect that the frequencies (kV_s) of the excited ion acoustic waves are greater than those (kV_d) of the low-frequency modes excited by the current convective instability for two reasons. One is that $V_d < V_s$. The other is that the current convective instability [Ossakow and Chaturvedi, 1979] and the ion acoustic instability [Kindel and Kennel, 1971] excite preferentially large-scale (i.e., small k) modes and short-scale (i.e., large k) modes, respectively. Thus, modes with frequencies less than the ion gyrofrequency (f_{ci}) were basically excited by the current convective instability in the VTF. As mentioned before, in the space experiments, the beam-plasma interactions occurred in a rather small region of the ionosphere. Hence, the beam-plasma interaction region in space experiments can be modelled as a uniform background plasma and, consequently, the current convective instability does not occur.

Comparison of our VTF experiments with space experiments have shown that we can simulate several features of the electron beam injection experiments in space plasmas. However, we can more closely simulate the naturally occurring plasma turbulence in the auroral ionosphere.

In conclusion, we have discussed some of our laboratory experiments with the Versatile Toroidal Facility (VTF) aimed at investigating space plasma turbulence. Our research results reported here have demonstrated that the VTF plasma device is capable of simulating some characteristic features of ionospheric plasma turbulence. Although the parameters of the VTF plasmas are quite different from those of space plasmas, our work has shown that laboratory experiments can adequately study some processes responsible for space plasma phenomena and complement the active experiments in space.

References

Dalkir, Y.R., M.C. Lee, and K.M. Groves, A mechanism responsible for the observation of symmetric lower hybrid sidebands and a low frequency mode in the upper ionosphere, *J. Geophys. Res.*, 97, 17195, 1992.

Ginet, G.P., and J. Ernstmeyer, VLF plasma waves generated by an electron beam in space, *Physics of Space Plasmas*, SPI Conference Proceedings and Reprint Series, Cambridge, Massachusetts, Number 11, 209, 1991.

Groves, K.M., M.C. Lee, and S.P. Kuo, Spectral broadening of VLF signals traversing the ionosphere, *J. Geophys. Res.*, 93, 14683, 1988.

Gurnett, D.A., et. al., Whistler mode radiation from the Spacelab 2 electron beam, *Geophys. Res. Lett.*, 13, 225, 1986.

Hui, B., et. al., Scattering of electron cyclotron resonance heating waves by density fluctuations in tokamak plasmas, *Nucl. Fusion*, 21, 339, 1981.

Inan, U.S., et. al., Modulated beam injection from the space shuttle during magnetic conjunctions of STS 3 with the DE 1 satellite, *Radio Sci.*, 19, 487, 1984.

Kellogg, P.J., S.J. Monson, W. Bernstein, and B.A. Whalen, Observations of waves generated by electron beams in the ionosphere, *J. Geophys. Res.*, 91, 12065, 1986.

Lee, M.C., and S.P. Kuo, Production of lower hybrid waves and field-aligned plasma density striations by whistlers, *J. Geophys. Res.*, 89, 10873, 1984.

Liao, C.P., J.P. Freidberg, and M.C. Lee, Explosive spread F caused by Lightning-induced electromagnetic effects, *J. Atmos. Terr. Phys.*, 51, 751, 1989.

Maggs, J.E., Coherent Generation of VLF Hiss, *J. Geophys. Res.*, 81, 1707, 1976.

Moriarty, D.T., *Electron cyclotron range of frequencies propagation in critically dense cold magnetoplasmas*, M.S. Thesis, M.I.T. Nuclear Engineering Department, 1992.

Ossakow, S.L., and P.K. Chaturvedi, Current convective instability in the diffuse aurora, *Geophys. Res. Lett.*, 6, 332, 1979.

Ott, E., Hui, B., and Chu, K.R., Theory of electron cyclotron resonance heating of tokamak plasmas, *Phys. Fluids*, 23, 1031, 1980.

Kindel, J.M., and C.F. Kennel, Topside current instabilities, *J. Geophys. Res.*, 76, 3055, 1971.

Reeves, G.D. et al., VLF wave emissions by pulsed and DC electron beams in space: 1. Spacelab 2 observations, *J. Geophys. Res.*, 93, 14699, 1988.

Winckler, J.R. et. al., Ion resonances and ELF wave production by an electron beam injected into the ionosphere: ECHO 6, *J. Geophys. Res.*, 89, 7565, 1984.

Winckler, J.R. et al., ECHO 7: An electron beam experiment in the magnetosphere, *EOS Trans. AGU*, 70, 657, 1989.

Winglee, R.M., and P.J. Kellogg, Electron beam injection during active experiments: 1. Electromagnetic wave emissions, *J. Geophys. Res.*, 95, 6167, 1990.

Yoo, C., *Plasma confinement optimization of the Versatile Toroidal Facility for ionospheric plasma simulation experiments*, M.S. Thesis, M.I.T. Nuclear Engineering Department, 1989.

PREPARATION AND CHARACTERIZATION
OF BLENDS OF ORGANIC POLYMERS
WITH NOVEL SILSESQUIOXANE MATERIALS

Charles J. Noel
Associate Professor
Department of Textiles and Clothing

The Ohio State University
1787 Neil Avenue
Columbus, OH 43210-1295

PREPARATION AND CHARACTERIZATION
OF BLENDS OF ORGANIC POLYMERS
WITH NOVEL SILSESQUIOXANE MATERIALS

Charles J. Noel
Associate Professor
Department of Textiles and Clothing
The Ohio State University

ABSTRACT

Polyhedral oligomeric silsesquioxane (POSS) materials are an important class of compounds having an oxygen:silicon ratio of 1.5, intermediate between that found in silicones and silica. Blends of three different POSS compounds, two monomeric (fully cyclohexylated T6 and fully cyclohexylated T8), and one a polymer [a poly(POSS)methacrylate], with two different organic polymers, poly(methyl methacrylate) (PMMA) and a thermoplastic elastomer based on a block copolymer of poly(tetramethylene ether) and nylon 12 (PEBAX 2533), were prepared by dissolving the materials to be blended in a common solvent and evaporating the solvent. The resulting blends were characterized by Thermogravimetric Analysis (TGA) and Differential Scanning Calorimetry (DSC). The blends with the monomeric POSS compounds, in which the POSS compounds were intended to serve as filler materials, appeared to phase separate more than the blends with the poly(POSS)methacrylate. Blends of poly(POSS) methacrylate with PMMA gave clear films at up to 6% loading of the POSS material, and clear films with PEBAX at up to 50% loading. The clear films with PEBAX whitened, however, upon stretching, suggesting that phase separation could be stress induced. The results of the TGA analyses showed that the solvent was never completely removed from blends containing PMMA regardless of the drying procedure employed, while solvent was always completely removed from the blends containing PEBAX. TGA traces of blends showed that there was little effect of POSS at low levels, but at higher levels traces became more like those of the POSS. Low levels of POSS had no effect on the T_g of PMMA and the T_g of the PEBAX was unaffected by any POSS at any level. Higher levels of POSS increased the T_g of PMMA by 5 to 10C. Low-temperature, low-energy endothermic transitions were observed in the DSC for the three POSS compounds; these were not observed in the glassy blends with PMMA, but were seen, shifted, in the rubbery blends with PEBAX.

PREPARATION AND CHARACTERIZATION
OF BLENDS OF ORGANIC POLYMERS
WITH POLYOLIGOSILSESQUIOXANE MATERIALS

Charles J. Noel

INTRODUCTION

The use of blending of polymeric materials to achieve a balance of properties not available in the unblended materials has become more and more important in polymer and material science. Current journals reporting work in these fields contain several articles in every issue reporting work involving polymer blends. Most of these works concern the blending of two or more different organic polymers. For example, Kyu, Park, and Cho (1) report that blends of polycarbonate (PC) and poly(phenyl methacrylate) (PPMA) can be prepared which are single phase in nature, and that, at a blend level of 12/88 PC/PPMA a non-birefringent material is obtained. (Both PC and PPMA are highly birefringent, but with opposite signs.) Thus blending produces a material with one property, birefringence, completely different from the single species. Recently Case and Honeycutt (2) published an article covering modelling techniques for studying polymer miscibility, based on thermodynamic principles. There is relatively little literature concerning the blending of organic polymers with polymers having inorganic character, such as preceramic polymers containing silicon. Asuka, Chian-Hua, and West (3) report on polymer blends of a polysilane (polysilastyrene) with polystyrene and polypropylene. Successful blending of the polysilane with polystyrene was achieved at the 5% polysilane level, but with polypropylene, some segregation of the polysilane was observed above the 1% level. The incorporation of polysilastyrene in polystyrene caused an increase in hardness, improved UV degradation, and decreased surface resistance, compared to unblended polystyrene. Similar but smaller effects were observed for the blends with polypropylene. Nago and Mizutani (4) report on the development of microporous polypropylene sheets containing polymethylsilsesquioxane (PMSS) filler. In this study, 42 parts of polypropylene were mixed with 56 parts PMSS and 2 parts additive/antioxidant. The polypropylene and PMSS did not blend; the spherical particles of PMSS were uniformly dispersed throughout the polypropylene, and provided the basis for development of micropores when the filled films were biaxially stretched. Blends of polysilastyrene and polycarbosilane with commercial thermoplastic elastomers were studied by Lichtenhan and Bolf (5). The silicon-containing polymers greatly enhanced the char-forming capability of the elastomers while increasing modulus and reducing stretch; the blends show potential as insulation materials in solid rocket motors. This study was undertaken to evaluate blends of polyhedral oligomeric silsesquioxane (POSS) materials with organic polymers. These

materials are hybrid organic/inorganic materials having an oxygen:silicon ratio of 1.5, intermediate between that found in silicones and silicates (6). The polymeric POSS material used in this study was a poly[3-T8(c-C_6H_{11})$_7$propyl methacrylate] having a number average molecular weight of 44,000. The T8 designation is shorthand for Si_8O_{12} arranged in a cube with a silicon atom at each corner and an oxygen atom between each silicon. Each silicon is thus bonded to three oxygen atoms; the fourth valency is a Si-C bond. Seven silicons are attached to cyclohehyl groups while the eighth is attached to the propyl methacrylate. This polymer will hereafter be referred to as P(POSS)MA. Blends were made of this polymer with both poly(methyl methacrylate) (PMMA) and a commercial thermoplastic elastomer which is a block copolymer of poly(tetramethylene ether) and nylon12 (PEBAX). These materials were chosen because both are fully amorphous polymers; one (PMMA) has a glass transition temperature of 115C and is a rigid plastic, while the other (PEBAX) has a glass transition temperature well below -80C and is an elastomer at room temperature. Blends were also made of the two organic polymers with monomeric POSS materials, to see if properties could be affected by the inclusion of these as fillers. The monmeric POSS materials used were fully cyclohexylated T8 and T6, [Si_8O_{12}(c-C_6H_{11})$_8$] and [Si_6O_9(c-C_6H_{11})$_6$], respectively. The objectives of the research were to determine the extent of blending possible between the POSS material and the organic polymers, and to determine the effects of blending on properties of the polymers. Only thermal properties were studied during this time period; mechanical properties will be evaluated in a later stage of this study.

METHODOLOGY

1. <u>Preparation of Blends</u>

Blends were prepared by dissolving weighed amounts of materials in an appropriate solvent, mixing thoroughly, and then allowing the solvent to evaporate slowly. Tetrahydrofuran (THF), chloroform, and toluene were used as solvents for the PMMA blends; only chloroform was used for the PEBAX blends. Normally, a total of 1.00g of solid material was dissolved in 20mL solvent in 25ml scintillation vials. Solution was achieved by a combination of techniques, which included shaking the vials for several hours using a wrist-action shaker at room temperature, exposing the shaken solutions to ultrasound at 45C for periods up to two hours, and heating at 60C for periods up to one hour. Heating was necessary for solutions of P(POSS)MA with PMMA in toluene, but was not necessary for the other solvents. Ultrasound was needed for all solutions containing PEBAX. In the absense of ultrasound the PEBAX swelled tremendously in the chloroform but did not dissolve; the ultrasound treatment broke the surface tension of the swollen particles, and

24-4

brought about solution of the polymer. All of the POSS materials dissolved quickly in all three solvents. Prior to pouring the solutions into aluminum weighing pans for solvent evaporation, the solutions were shaken vigorously to insure proper mixing. All solutions were clear and visually homogeneous, with no optical evidence of concentration or density gradients. The solutions were poured into alumonum weighing pans, and a 400mL beaker was inverted over each pan, to allow the evaporation to proceed slowly. Solid materials were obtained after overnight evaporation in most cases. Following evaporation, the blends were usually dried in a vacuum oven at 75C and -20in Hg overnight to remove residual solvent.

2.   Thermal Analysis

The materials prepared were analyzed by both Thermogravimetric Analysis (TGA) and Differential Scanning Calorimetry (DSC). TGA provides a means for following sample weight loss as temperature is increased at a constant rate. The instrument used was a T.A. Instruments Model 951 equipped with a Model 2000 Thermal Analyst Controller. Samples weighing 15-20mg were placed on a platinum boat suspended on one arm of a microbalance, and the temperature was raised from room temperature to 1000C at a rate of 10C/m under an atmosphere of nitrogen gas flowing at 50mL/m. Weight-temperature traces were plotted and analyzed. DSC provides a means for observing first-order thermal transitions such as melting and crystallization as well as second-order thermal transitions such as glass transition temperature (T_g). The latter property was of primary interest here, since it was expected that the T_g values for the polymer/polymer blends would increase with concentration of the P(POSS)MA as it should have a higher glass transition temperature than either of the two organic polymers. The instrument was a T.A. Instruments Model 912 also equipped with a Model 2000 Thermal Analyst Controller. Samples weighing 5-15 mg were placed in an aluminum pan (weighing approximately 12-13 mg) and an aluminum lid (9-10mg) was crimped onto the pan to cover the sample. An empty aluminum pan/lid was used as reference. Hermetically sealed pans were not used. Since TGA showed that solvent was never completely removed for the blends with PMMA, and was always completely removed for the blends with PEBAX, different temperature profiles were used for the two sets of blends. For the PMMA blends, the temperature was first raised from room temperature to 180C at the rate of 10C/m, the sample was cooled back to room temperature, and the same sample was then reheated from room temperature to 150C, also at 10C/m. The first heating allowed residual solvent to evaporate from the material, and the second heating showed the T_g transition unequivocally. Since there was no residual solvent in the PEBAX blends and since the T_g for this elastomeric material is well below room temperature, only one heating was done for these materials. Samples were cooled to -150C with liquid nitrogen, then

the temperature was raised from -150C to +150C at the rate of 10C/m. All DSC runs were made with an atmosphere of nitrogen gas in the DSC cell. Both TGA and DSC traces were obtained for all the unblended starting materials as well as for all prepared blends.

RESULTS

1. Blends Prepared   The composition of the blends prepared is given in Table I. Blends of PMMA with the three POSS materials were prepared at the 3, 6, 9, 12, and 15% (w/w) levels of the POSS compounds. Additional blends (or mixtures) with higher levels of POSS were prepared for the T6 and P(POSS)MA. Clear films were obtained at only the 3 and 6% levels of POSS material. Blends containing 9 - 15% POSS tended to be opaque, and at higher POSS levels, white powdery material (presumably POSS) was transferred from the blends to the fingers when the films were handled. It was noted that, when the PMMA and the POSS materials were dissolved in solvent, the POSS material always dissolved much more rapidly than the PMMA. It is hypothesized that, during solvent evaporation and film formation, the PMMA may precipitate first, leaving the solution relatively rich in the POSS compound. In the final stages of evaporation the POSS material precipitates at the surface of the film producing a phase-separated system. This solvent-induced phase separation could be eliminated if blending were done in the melt, in the absense of solvent. This could not be done with the equipment at hand when this study was carried out, but is recommended for future work. Blends of PEBAX-2533 with the three POSS materials were prepared at the 17, 34, 50, 67, and 83% (w/w) levels of the POSS compounds. Unlike the PMMA blends, these films were fairly clear up to 50% POSS and opaque at higher POSS levels. All of these elastic films, however, tended to whiten when stretched, suggesting stress-induced phase separation. After stretching, white powdery material was transferred from high POSS films to fingers when these films were handled.

2. Thermogravimetric Analysis (TGA)   TGA traces for PMMA (as received) show a single smooth decomposition step going from 0 to virtually 100% weight loss; this is consistent with the known thermal degradation behavior of PMMA which depolymerizes to monomer with over 99% yield and little discernible residue. Traces for the PMMA cast from solvent showed a two step weight loss, with residual solvent coming off first, followed by the complete depolymerization of the remaining polymer. The amount of solvent retained by the cast films of PMMA ranged from 22.4 to 5.61% by weight, depending upon the solvent used and the extent and type of drying done. (The solvent-cast PMMA films were never dryed above the T_g of PMMA, to prevent bubbling of the films.) Initiation of depolymerization of PMMA occurred near 360C, the inflection point of the weight loss curve occurred near 390C and the process was complete near

24-6

TABLE I

Composition of Blends Prepared of POSS Compounds With PMMA and PEBAX

PMMA/POSS	T6 POSS	T8 POSS	P(POSS)M
100/0	x	x	x
97/3	x	x	x
94/6	x	x	x
91/9	x	x	x
88/12	x	x	x
85/15	x	x	x
80/20	x		
75/25			x
70/30	x		
60/40	x		
50/50	x		x
25/75	x		x
13/87	x		
0/100			x
PEBAX/POSS			
100/0	x	x	x
85/15	x		
83/17	x	x	x
66/34	x	x	x
50/50	x	x	x
33/67	x	x	x
17/83	x	x	x
0/100			x

410C. (Initiation and completion temperatures were determined as the intercepts of the tangent to the weight loss curve at the inflection point, extrapolated to the baselines preceding and following the process.) The TGA traces for PEBAX, both as received and as cast from chloroform, showed a single step weight loss with an onset near 390C, an inflection point near 425C, and completion near 460C. A residue of around 1.5% was indicated for this material. The decomposition mechanism for the PEBAX is presumably random chain scission. TGA traces for the monomeric T6 POSS compound showed a two step weight loss curve. The bulk of the weight loss (~88%) occurs in a step initiating near 344C with an inflection near 377C and ending near 390C. Further gradual weight loss (~5%) continues until near 535C, leaving a residue of about 7%, through 1000C. This material is known to sublime and this sublimation is accounted for in the first major weight loss step. The material was observed to condense on the walls of the glass chamber outside the furnace of the TGA apparatus during this weight loss step. The momomeric T8 POSS material is also known to sublime, but at higher temperatures. TGA traces for this material showed a single step weight loss curve with an onset temperature near 440C, an inflection near 480C and completion near 490C. Weight loss in this step was 98% with a residue of 2%. The TGA traces for the polymeric POSS compound, P(POSS)MA, were very different. Most of the weight loss (~52%) occurred in a step with an extrapolated onset temperature near 360C, an inflection near 460C and completion near 525C. There was a continued gradual weight loss of another 5% between 525C and around 700C, then a stable baseline indicating ~43% residue through 1000C. The initial weight loss step appeared to begin much earlier than the extrapolated onset temperature of near 360C; definite deviations from the initial baseline were noted at temperatures as low as 250C. Being a substituted poly(n-propyl-methacrylate), the expected decomposition mechanism is depolymerization to monomer; this may be altered by the presence of the $T8(c-C_6H_{11})_7$ group which makes the monomer molecular weight 1124. If depolymerization does occur, the monomer, being unsymmetrical, apparently does not sublime as does the fully cyclohexylated T8, and the lack of sublimation could lead to the considerable residue formation as the monomer itself thermally decomposes.

The TGA traces for the blends of PMMA with T6, following the solvent weight loss step, were virtually identical with the traces for PMMA alone, up to 15% T6. At higher levels, the decomposition onset, inflection, and completion temperatures are shifted to lower temperatures, a second decomposition step is observed which becomes larger at the higher T6 levels (73 and 87%) and residues of 18 and 24% are observed at these T6 levels. Since the T6 itself produced a residue of only about 7%, the presence of 13 - 27% PMMA apparently prevented the sublimation of much of the T6, leading to the much higher residue formation. Blends of PMMA with T8 were only prepared up to the 15% T8 level. The TGA traces for these blends, following

24-8

the solvent loss, were also virtually identical to those for unblended PMMA, as were the T6 blends at these levels. The TGA traces for the blends of PMMA with P(POSS)MA were similar to the traces for PMMA alone up to 15% P(POSS)MA levels, except that residue increased slightly with increasing P(POSS)MA content, from near 0% for PMMA up to ~3% for 15% P(POSS)MA. At 25% P(POSS)MA the second decomposition step associated with this material was observed, accounting for ~6% weight loss, and the residue was ~5%. At 50% P(POSS)MA, the second secomposition step was slightly larger (~7% weight loss), but the residue was only ~3%, and at 75% P(POSS)MA the second decomposition step was much larger (~12% weight loss) and the residue was ~10%. This residue level is considerably less than expected, since P(POSS)MA by itself yields ~43% residue, a composition containing 75% P(POSS)MA should leave a residue of ~33%. This result is completely different than that obtained with high levels of the T6 POSS compound, where higher residues than expected were obtained.

The TGA traces for the blends of PEBAX with the POSS compounds were different for each compound. For the blends with monomeric T6, the weight loss curves were all single step, with the exception of the 34/66 blend, which had a small second step. Residues were all small (~1 - 3%) except for the 34/66 blend, which had a residue of ~9%. The temperatures associated with the main weight loss step, onset, inflection, and completion, tended to decrease with increasing T6 level, except that the temperatures for the 17/83 blend were higher than those for the 34/66 blend. For the blends with monomeric T8, single step weight loss curves were obtained over the entire range studied, and residues remained in the 1-2% range. There was a steady increase in associated onset, inflection, and completion temperatures with increasing T8 content. For the blends with P(POSS)MA, the TGA traces resembled those for PEBAX alone up to the 34/66 blend level, except that residue increased with increasing P(POSS)MA content, from ~5% for the 83/17 blend to ~15% for the 34/66 blend. At the 17/83 blend level, the TGA trace resembled that for pure P(POSS)MA, with a residue of ~40%.

A summary of TGA results appears in Table II. Representative TGA curves are given in Figures 1A - 1C.

3. Differential Scanning Calorimetry (DSC)   DSC traces were used to evaluate the glass transition temperatures of the polymeric starting materials and their blends. DSC traces were also obtained for the monomeric T6 and T8 POSS compounds. For PMMA, both as received and as cast from either chloroform or from THF, glass transition temperatures of ~105 - 110C were observed; these values are consistent with the literature. However, for PMMA cast from toluene, values of ~95 - 97C were obtained. Since the T_g values were obtained on second heatings of samples following an initial heating to bring about solvent removal, the reason for

TABLE II

Summary of TGA Data for POSS Blends with PMMA and PEBAX

Material		T(onset)	T(infl)	T(comp)	Residue	T(onset)	T(infl)	T(comp)	Residue
PMMA		360	390	410	0.0				
PEBAX		390	425	460	1.5				
T6 POSS		345	377	390		510	535	555	7.0
T8 POSS		440	480	492	2.1				
P(POSS)MA	(250)	36	460	520	43.1				
PMMA/T6	97/3	358	385	406	0.4				
	94/6	352	385	407	0.5				
	91/9	354	386	409	0.7				
	88/12	358	382	405	0.1				
	85/15	357	382	409	0.9				
	80/20	355	382	409	4.4				
	70/30	352	380	408	2.9				
	60/40	348	379	310	2.8				
	50/50	344	376	409		505	540	560	4.2
	27/73	340	369	400		492	525	550	17.8
	13/87	318	353	393		492	524	550	23.7
PMMA/T8	97/3	362	393	410	0.2				
	94/6	366	397	415	1.2				
	91/9	362	391	416	1.1				
	88/12	351	387	415	2.6				
	85/15	362	390	418	1.1				
PMMA/P(POSS)MA									
	97/3	360	386	410	0.6				
	94/6	362	389	412	1.5				
	91/9	359	384	408	1.8				
	88/12	361	387	411	1.8				
	85/15	362	390	418	3.3				
	80/20	363	389	412	1.2				
	75/25	357	384	413		510	544	552	5.0
	50/50	353	381	409		502	530	545	2.5
	25/75	353	378	409		510	531	552	9.7
PBAX/T6	85/15	386	417	453	1.4				
	83/17	391	431	461	1.3				
	66/34	378	431	462	3.1				
	50/50	364	431	463	2.6				
	34/66	332	385	433	9.2				
	17/83	354	393	420	2.3				
PBAX/T8	83/17	402	435	458	1.5				
	66/34	404	441	465	1.5				
	50/50	410	454	478	1.4				
	34/66	425	471	482	1.7				
	17/83	431	473	486	2.1				
PEBAX/P(POSS)MA									
	83/17	392	423	455	4.6				
	66/34	395	428	460	6.0				
	50/50	383	420	459	9.1				
	34/66	386	426	467	14.8				
	17/83	(270) 37	467	505	39.6				

these low values is not known. For PEBAX, the glass transition temperature was found to be ~-76C (197K). For P(POSS)MA, no glass transition temperature was observed. For the monomeric T6 POSS, there was no glass transition temperature, but a sharp endotherm was observed near 50C, having an energy of ~9.7 J/g. This was reproducible when the sample was cooled and reheated. For the monomeric T8 POSS, no transitions were observed above room temperature, but a double endotherm (two peaks) was observed with peak temperatures near -31C and -25C. The total energy of the two peaks was ~13.8 J/g. For these two materials, these transitions may be related to the conformations of their cyclohexyl groups. An additional sample of P(POSS)MA cast from chloroform was scanned first from room temperature to 150C to remove solvent, then from -150C to 500C. As before, no glass transition was observed. An endotherm similar to those observed with T6 and T8 was observed near -3C, with a transition energy of ~0.7 J/g. A marked decomposition exotherm was observed with an onset temperature of ~210C, consistent with the TGA result that weight loss began to occur gradually near 250C. The decomposition exotherm first peaked near 235C, dipped, then rose to a maximum near 330C. The total energy associated with this broad exotherm was ~375 J/g.

Blends of PMMA with T6 POSS gave T_g values near 116 to 117C, once a nominal blend level of about 9% T6 was realized. Blends of PMMA with T8 POSS gave T_g values of 113 to 117C for nominal levels of T8 of 9 to 15%. Blends of PMMA with P(POSS)MA gave T_g values ranging from 105 to 110C for nominal levels of P(POSS)MA of 3 to 15% and T_g values of 116 to 118C for nominal levels of 25 to 75% P(POSS)MA (when cast from chloroform). Three 80/20 PMMA/P(POSS)MA blends cast from toluene gave T_g values of 94 to 95C. These values are the same as for the three PMMA films cast from toluene; no reason for the lower values are known.

Blends of PEBAX with POSS compounds all showed the glass transition temperature of the PEBAX (~-71 to -76C) regardless of the nature of the POSS material used or the blend level. The DSC traces also showed the endotherms associated with the POSS compounds, but the peak temperatures were shifted, indicating some interaction with the PEBAX. Blends of PEBAX with P(POSS)MA showed a single endotherm near 10C, about 13C higher than for the P(POSS)MA alone. Blends of PEBAX with T8 POSS showed two endotherms, one near -25C with its energy increasing with increasing T8 content, and the other near +25C with its energy tending to decrease with increasing T8 content. Blends of PEBAX with T6 POSS showed two endotherms, one near 5C and the second near 50C. Peak energies showed no particular pattern with increasing T6 content. The fact that fully substituted POSS compounds exhibit sharp low temperature endotherms and that the temperatures of these shift when blended with polymers having no transitions in these temperature regions should be studied further.

A summary of the DSC results appears in Table III. Representative DSC curves are given in Figure 2A - 2C.

SUMMARY AND CONCLUSIONS

Blends of three POSS compounds with both PMMA and PEBAX were prepared over a wide range of nominal blend levels (Table I) by casting from solution, usually with chloroform as solvent. Blends of these materials with PMMA were in the form of glassy films, and clear, transparent films were only obtained at low levels (3 to 6%) of POSS material. At these low levels, the TGA and DSC traces of the blends were similar to those of unblended PMMA. At higher levels of POSS material the films were opaque, indicating phase separation, and the glass transition temperatures of the blended materials tended to be ~5 to 10C higher than those for PMMA and for the low level blends. There was no clear cut relationship to either the nature of the POSS material or the blend level, once 9 to 15% POSS was present. At POSS levels up to ~15%, TGA traces tended to be similar to those for PMMA alone, except for slight increases in residue. Higher levels of T6 or P(POSS)MA in blends with PMMA tended to give TGA curves which shifted toward those of the POSS compound, except that residue formation was enhanced in the T6 blends and reduced in the P(POSS)MA blends. Blends of PMMA having greater than 15% T8 were not prepared.

Blends of the three POSS compounds with PEBAX produced fairly clear elastic films up to the 50% level of the POSS, and opaque elastic films at higher POSS levels. The stretchiness of the films appeared to decrease with increasing POSS levels. The clear films whitened on stretching, suggesting stress-induced phase separation. TGA traces showed that the blends exhibited degradation behavior intermediate between that of the unblended PEBAX and the unblended POSS, depending on the POSS level. DSC traces showed no effect on the glass transition temperature of the PEBAX. DSC traces of the unblended POSS compounds revealed the presence of low-temperature, low-energy endotherms for each of these materials. These may be associated with the conformation of the cyclohexyl groups on these compounds. These endotherms shifted in both temperature and size with POSS level in the rubbery PEBAX blends, and were not observed in the rigid, glassy blends with PMMA. Some interaction of the POSS material with the PEBAX is apparently taking place.

Blends of POSS materials with PMMA were never free of solvent, and solvent may have contributed to phase separation by retaining POSS material in solution longer than PMMA during the evaporation phase, producing concentration gradients, and leaving the film surface rich in POSS material. This might be minimized by either mixing in the melt or by coextruding the materials, but care would have to be taken to avoid thermal degradation of P(POSS)MA.

24-12

TABLE III

Summary of DSC Data for POSS Blends with PMMA and Pebax

	PMMA Blends				**PEBAX Blends**		
POSS	Blend	Solvent	TG	POSS	Blend	Solvent	TG
none	100/0	none	110	none	100/0	none	-76
	100/0	THF	110	none	100/0	Chloroform	-77
	100/0	Chloroform	109				
	100/0	Toluene	96				
T6	97/3	Chloroform	103	T6	83/17	Chloroform	-76
	94/6	"	107		66/34	"	-75
	91/9	"	117		50/50	"	-76
	88/12	"	110		34/66	"	-76
	86/14	THF	115		17/83	"	-76
	86/14	Chloroform	111				
	85/15	"	116				
	80/20	"	117				
	70/30	"	117				
	60/40	"	117				
	50/50	"	117				
	27/73	"	116				
	13/87	"	116				
T8	97/3	Chloroform	106	T8	83/17	Chloroform	-76
	94/6	"	106		66/34	"	-76
	91/9	"	116		50/50	"	-71
	88/12	"	115		34/66	"	-74
	85/15	"	115		17/83	"	-76
P(POSS)MA	97/3	Chloroform	105	P(POSS)MA	83/17	Chloroform	-76
	94/6	"	105		66/34	"	-76
	91/9	"	108		50/50	"	-76
	88/12	"	111		34/66	"	-76
	85/15	"	107		17/83	"	-76
	80/20	Toluene	95		0/100	"	(none)
	75/25	Chloroform	116				
	50/50	"	118				
	25/75	"	117				
	0/100	"	(none)				

REFERENCES

1.   Kyu, T., Park, D. and Cho, W., *J. Applied Polymer Science* 44, 2233-37 (1992)

2.   Case, F.H., and Honeycutt, J.D., *Trends in Polymer Science* 2 (No. 8), 259-66 (1994)

3.   Asuke, T., Chien-Hua, Y., and West, R., *Polymer Blends of a Polysilane with Polystyrene and Polypropylene,* preprint, to be published in *Macromolecules* (1994)

4.   Nago, S. and Mizutani, Y., *J. Applied Polymer Science* 50, 1815-22 (1993)

5.   Lichtenhan, J.D. and Bolf, A.B., *unpublished* (1993)

6.   Lichtenhan, J.D., *Silsesquioxane-Based Polymers,* preprint, to be published in the *CRC Encyclopedia of Polymeric Materials* (1994)

FIGURE 1A

Sample: PMMA, AS RECEIVED
Size: 28.5790 mg
Method: POLYMERS
Comment: RATE 10°C/MIN, GN2 ATM 50 ML/MIN

TGA

File: C: TGANOEL.001
Operator: CJNOEL
Run Date: 23-Jun-94 13:34

363.99°C

99.63 %
389.35°C (I) (28.47 mg)

408.68°C

Weight (%)

Temperature (°C)

TGA V5.1A DuPont 2000
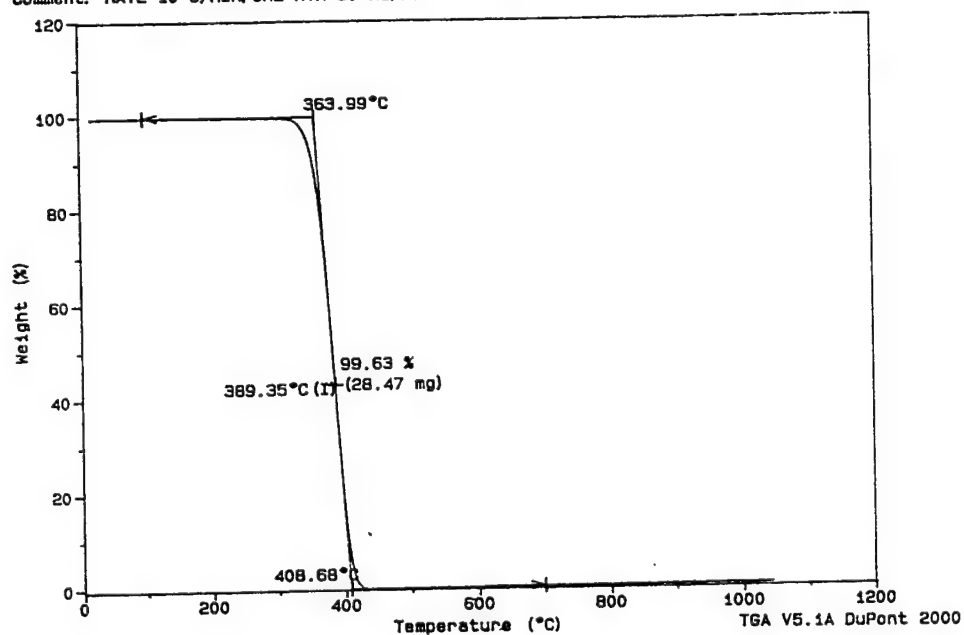
Sample: CJN-1-37-3 PEBAX FR CHL, VAC-DR
Size: 26.8300 mg
Method: POLYMERS
Comment: RATE 10°C/MIN, GN2 ATM 50 ML/MIN

TGA

File: C: TGACNOEL.009
Operator: C J NOEL
Run Date: 15-Jul-94 07:20

393.41°C

97.12 %
(26.06 mg)

426.71°C (T)

Residue:
1.862 % RESIDUE
(0.4995 mg)

460.01°C

Weight (%)

Temperature (°C)

TGA V5.1A DuPont 2000

24-15
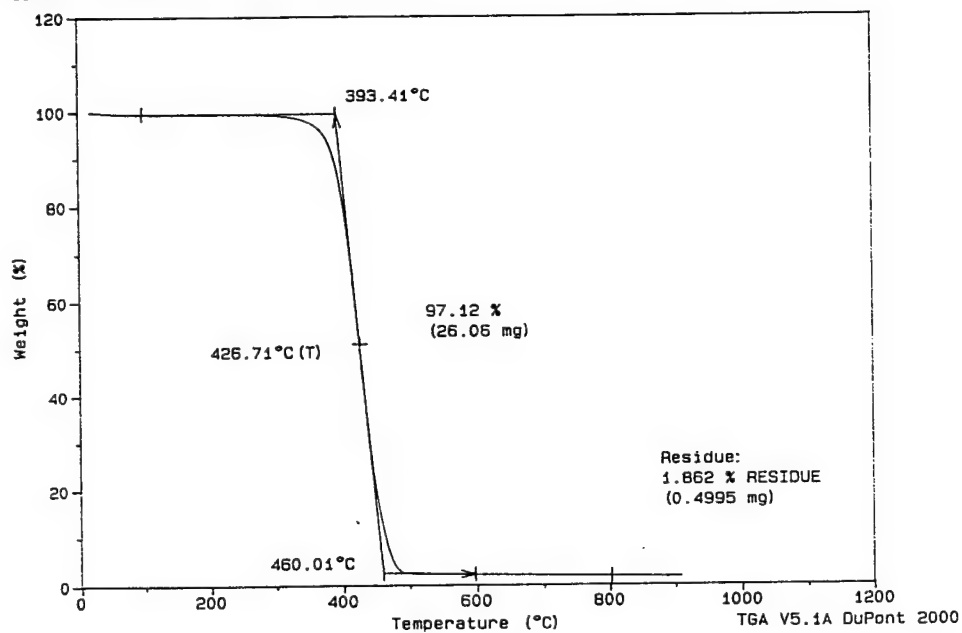
FIGURE 1B

Sample: T6CY6, PURE MATERIAL
Size: 14.9300 mg
Method: POLYMERS
Comment: RATE 10°C/MIN, GN2 ATM 50 ML/MIN

TGA

File: C: TGANOEL.003
Operator: CJNOEL
Run Date: 24-Jun-94 08: 21

344.21°C

88.45 %
(13.21 mg)

377.23°C (I)

4.736 %
(0.7071 mg)

509.56°C

389.92°C

533.04°C (I)

554.07°C

TGA V5.1A DuPont 2000
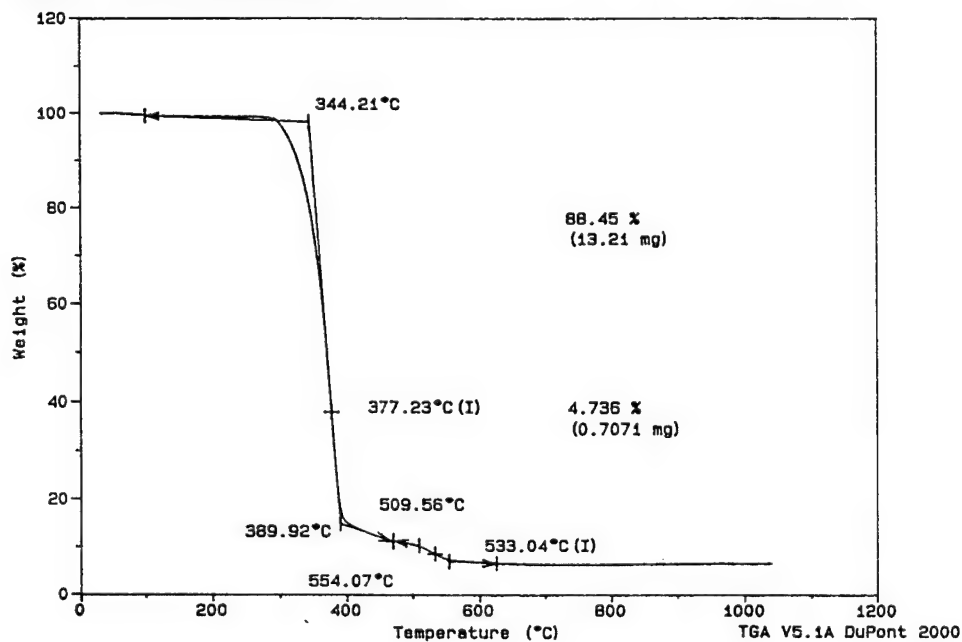
Sample: PURE T6CY8 AS REC'D
Size: 19.0620 mg
Method: POLYMERS
Comment: RATE 10°C/MIN, GN2 ATM 50 ML/MIN

TGA

File: C: TGACNOEL.054
Operator: C J NOEL
Run Date: 4-Aug-94 15: 06

442.73°C

97.83 %
(18.65 mg)

479.05°C (I)

Residue:
2.234 % RES AT 800C
(0.4258 mg)

492.46°C

TGA V5.1A DuPont 2000

24-16
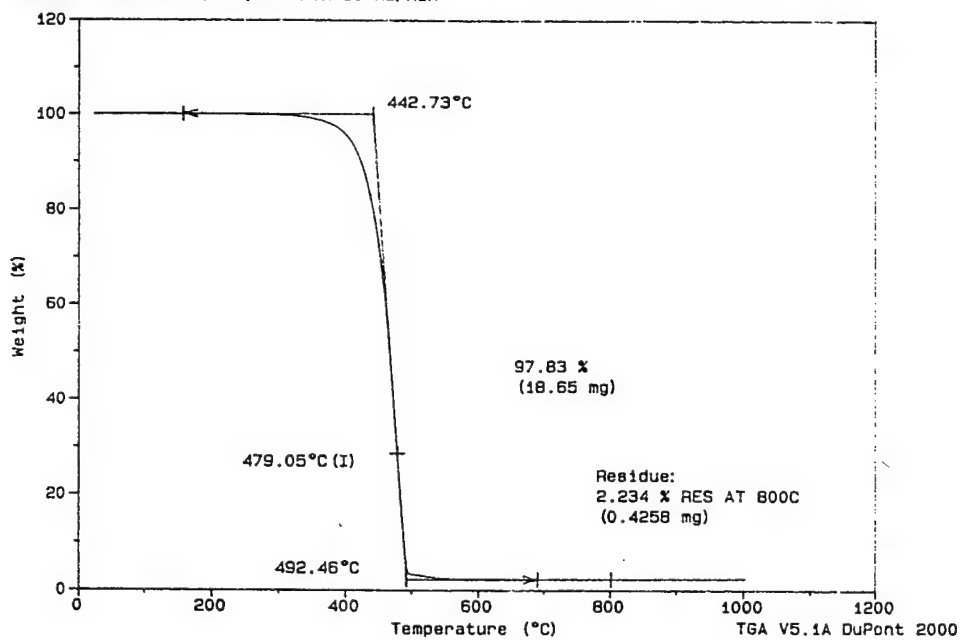
FIGURE 1C

Sample: P(POSS)MA YO IV-33 AS RECEIVED   TGA   File: C: TGACNOEL.017
Size: 16.0620 mg   Operator: C J NOEL
Method: POLYMERS   Run Date: 19-Jul-94 12:13
Comment: RATE 10°C/MIN, GN2 ATM 50 ML/MIN
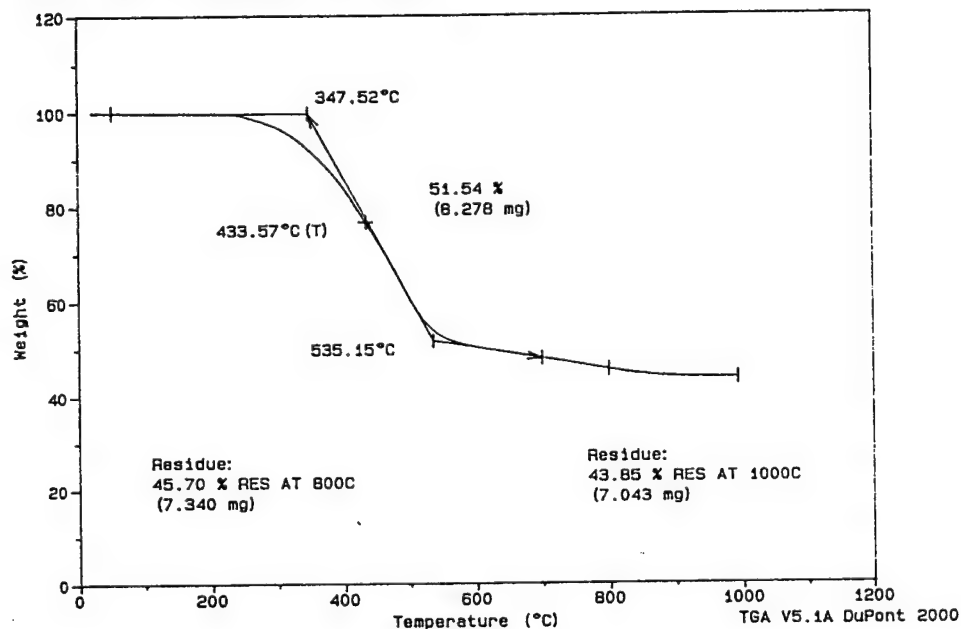
Sample: CJN-1-38-4 PMMA/PPMA 25/75   TGA   File: C: TGACNOEL.015
Size: 16.1620 mg   Operator: C J NOEL
Method: POLYMERS   Run Date: 19-Jul-94 07:19
Comment: RATE 10°C/MIN, GN2 ATM 50 ML/MIN

24-17
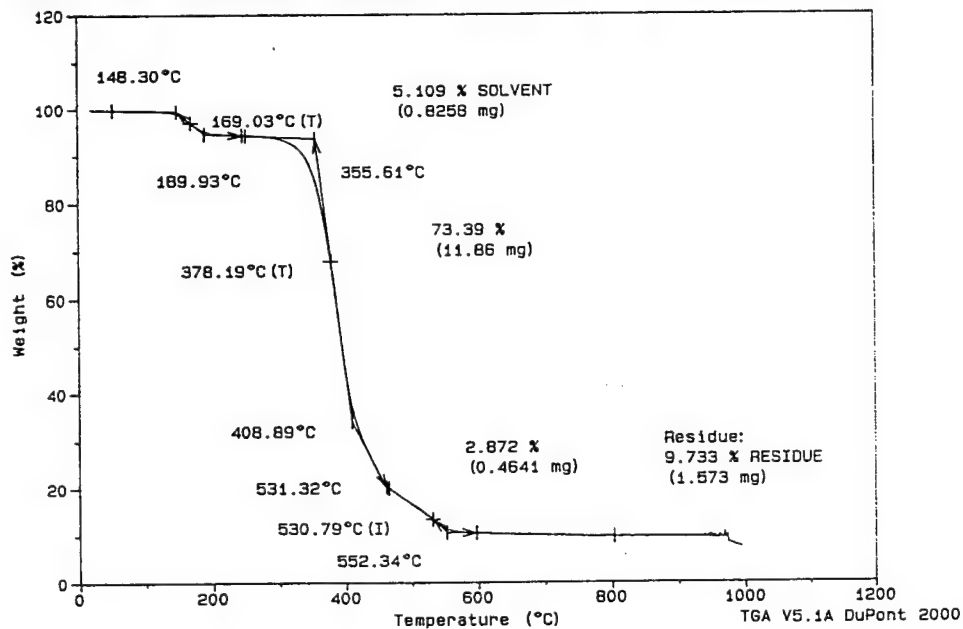
FIGURE 2A

Sample: CJN-1-36-1 PMMA 2ND HTG   DSC   File: DSCCNOEL.18
Size: 4.5900 mg   Operator: CHARLES J. NOEL
Method: CJN-DSC-TG/BLENDS   Run Date: 26-Aug-94 15: 29
Comment: RATE 10°C/MIN, GN2 ATM, AL PANS

99.97°C

110.52°C (I)

114.29°C

DSC V4.0B DuPont 2000
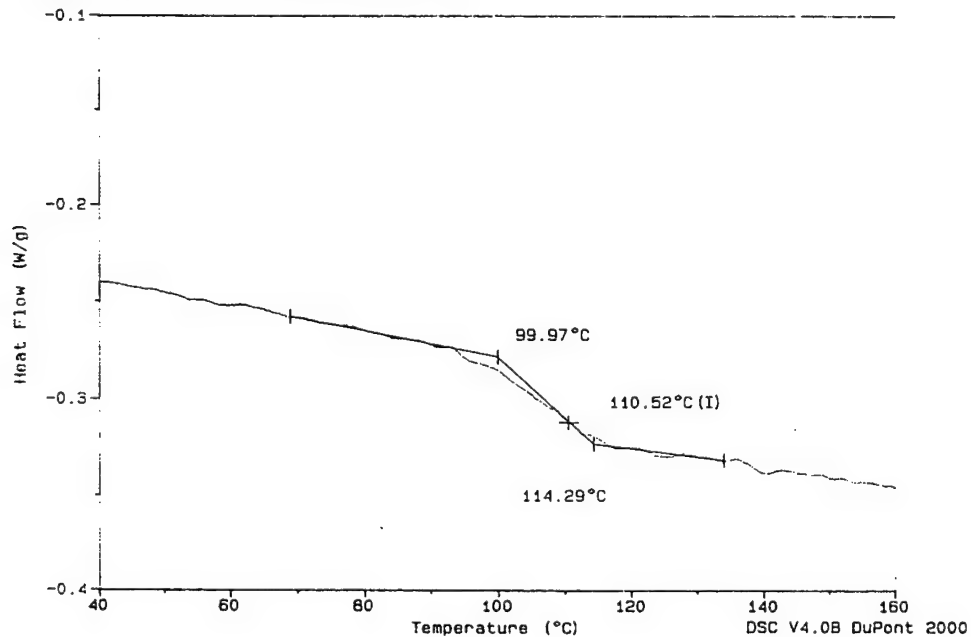
Sample: CJN-1-37-3 #4 PEBAX CAST HT#1   DSC   File: DSCCNOEL.94
Size: 5.2300 mg   Operator: CHARLES J. NOEL
Method: CJN BLENDS W PEBAX   Run Date: 2-Sep-94 14: 36
Comment: RATE 10°C/MIN, GN2 ATM, AL PANS

-80.91°C

-76.31°C (I)

-70.99°C

DSC V4.0B DuPont 2000

24-18
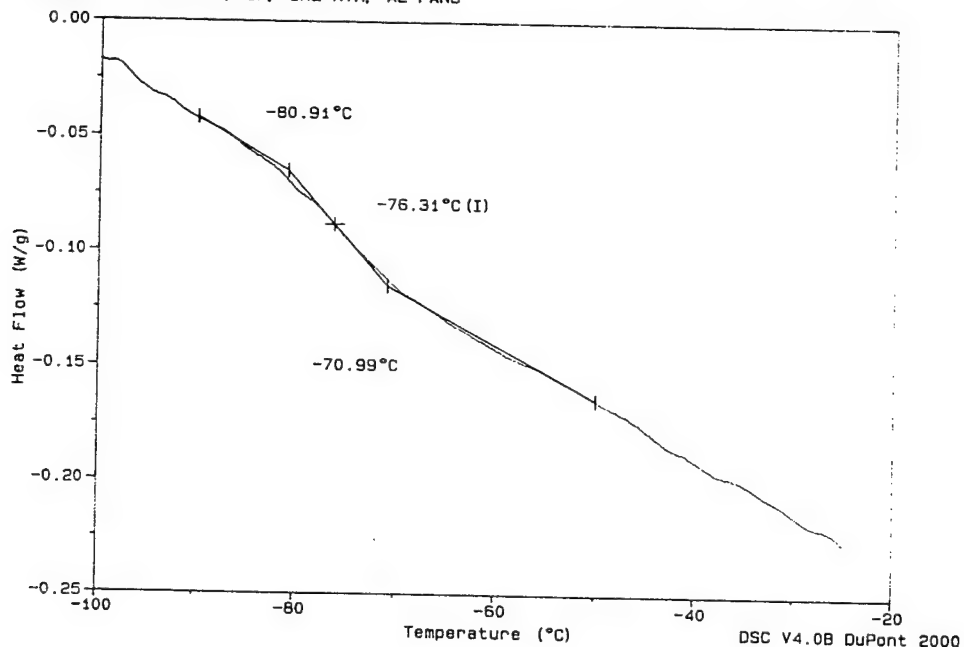
FIGURE 2B

Sample: CJN-1-35-4 PMMA/T6 2ND HTG
Size: 8.5900 mg
Method: CJN-DSC-TG/BLENDS
Comment: RATE 10°C/MIN, GN2 ATM, AL PANS

DSC

File: DSCCNOEL.10
Operator: CHARLES J. NOEL
Run Date: 19-Aug-94 09:45

105.08°C

110.67°C(I)

116.54°C

Heat Flow (W/g)

Temperature (°C)

DSC V4.0B DuPont 2000
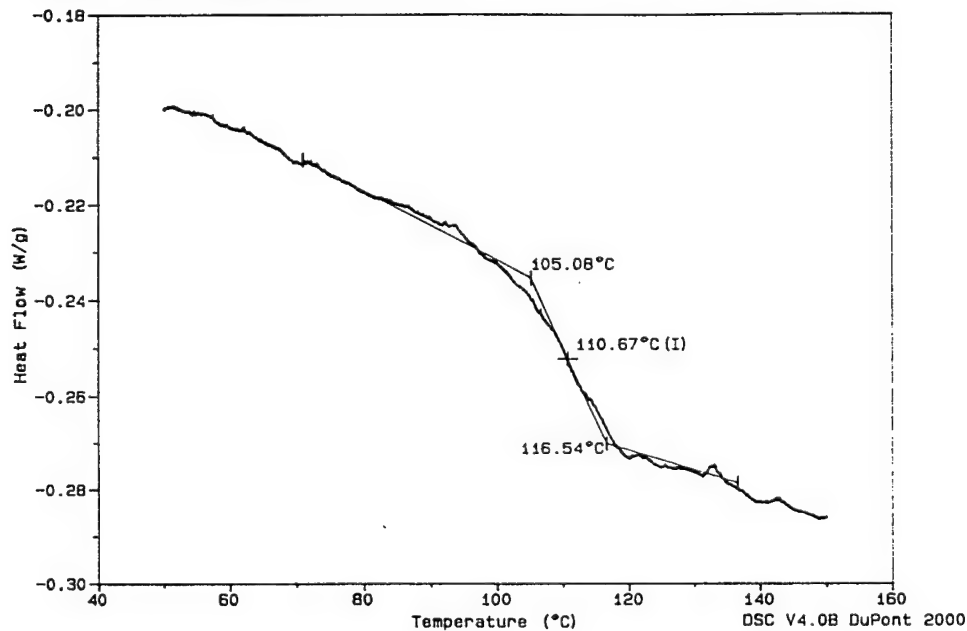
Sample: CJN-1-38-3 PMMA/PPMA 49/51 HT#2
Size: 4.7400 mg
Method: CJN-DSC-TG/BLENDS
Comment: RATE 10°C/MIN, GN2 ATM, AL PANS

DSC

File: DSCCNOEL.64
Operator: CHARLES J. NOEL
Run Date: 31-Aug-94 15:52

111.12°C

117.61°C(I)

120.57°C

Heat Flow (W/g)

Temperature (°C)

DSC V4.0B DuPont 2000

24-19
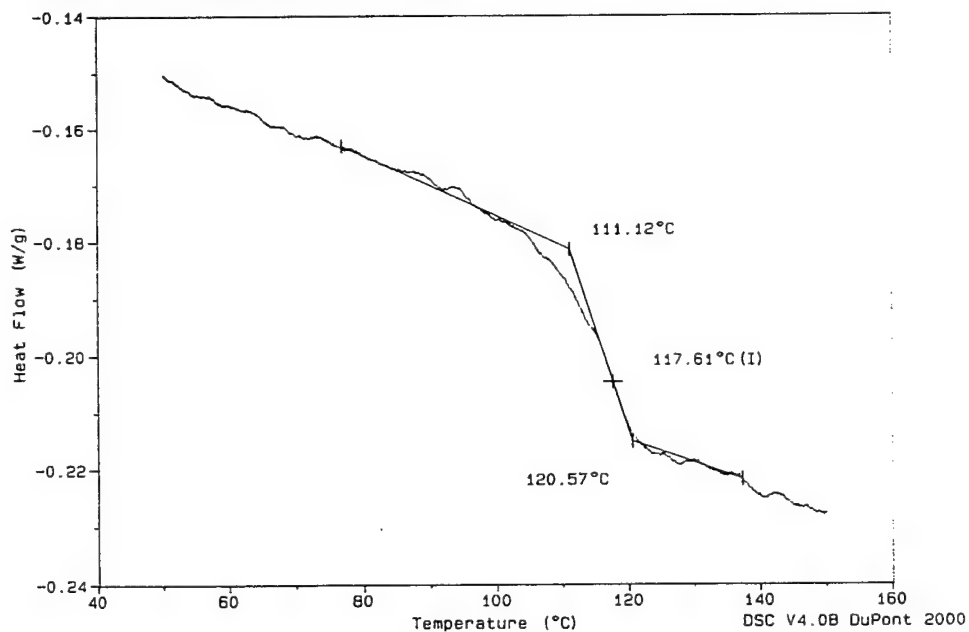
FIGURE 2C

Sample: T6CY6-AS RECD 1ST HTG
Size: 6.5000 mg
Method: CJN-DSC-TG/BLENDS
Comment: RATE 10°C/MIN, GN2 ATM, AL PANS

DSC

File: DSCCNOEL.11
Operator: CHARLES J. NOEL
Run Date: 19-Aug-94 10:22

52.20°C
9.526J/g

54.86°C

Heat Flow (W/g)
Temperature (°C)
DSC V4.0B DuPont 2000
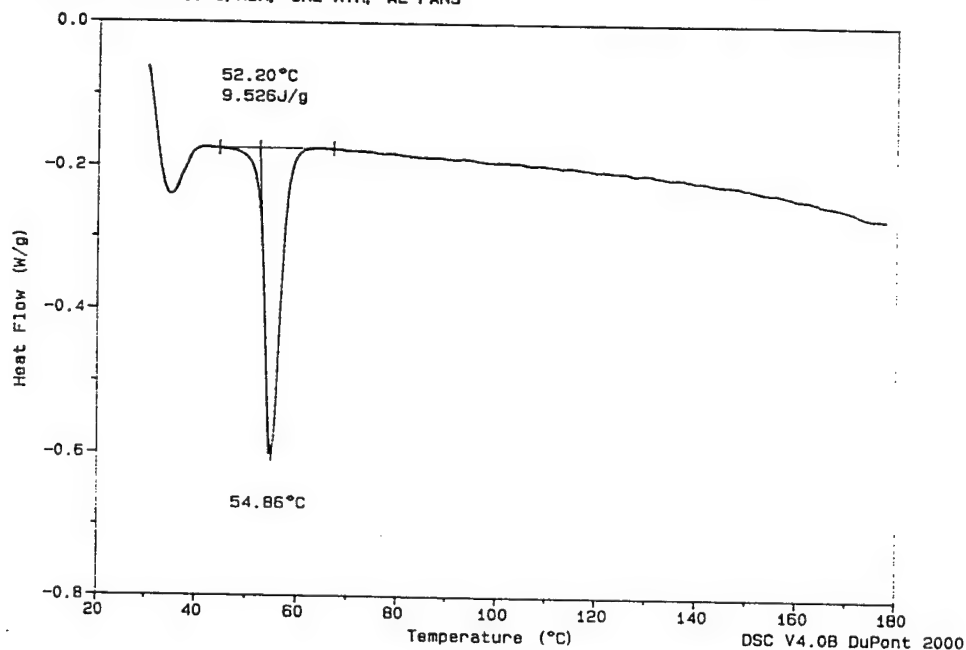
Sample: T6Cy8 AS RECEIVED, 2ND HT
Size: 9.6200 mg
Method: CJN BLENDS W PEBAX
Comment: RATE 10°C/MIN, GN2 ATM, AL PANS

DSC

File: DSCCNOEL.113
Operator: CHARLES J. NOEL
Run Date: 7-Sep-94 15:48
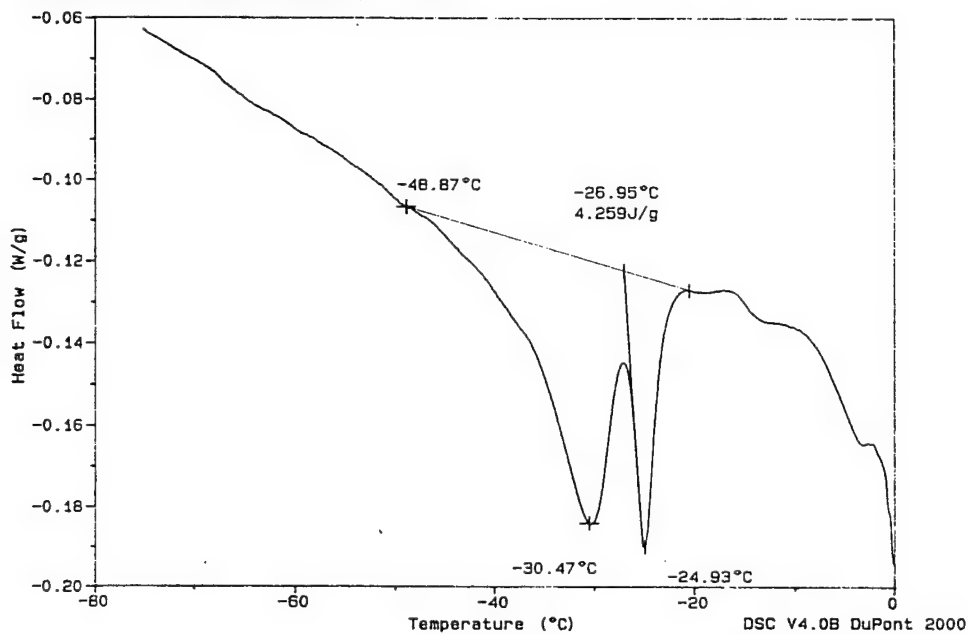
-48.87°C
-26.95°C
4.259J/g

-30.47°C
-24.93°C

Heat Flow (W/g)
Temperature (°C)
DSC V4.0B DuPont 2000

24-20

A HYBRID ALGEBRAIC EQUATIONS OF MOTION - NEURAL ESTIMATOR
TO IMPLEMENT THE DIRECT CONTROL METHOD

Hayrani Öz
Associate Professor
Department of Aerospace Engineering, Applied Mechanics & Aviation

The Ohio State University
Room 328 Bolz Hall
2036 Neil Ave
Columbus, OH 43210

A HYBRID ALGEBRAIC EQUATIONS OF MOTION - NEURAL ESTIMATOR TO IMPLEMENT THE DIRECT CONTROL METHOD

Hayrani Öz
Associate Professor
Department of Aerospace Engineering, Applied Mechanics & Aviation
The Ohio State University

Abstract

The study of dynamic systems without resorting to or any knowledge of differential equations is known as the "Direct Method". In this method, algebraic equations of motion describe the system. By this approach algebraic optimal control laws can be derived in explicit form for general nonlinear time-variant and time-invariant systems by minimizing an algebraic performance measure. The essence of the approach is based on using assumed-time-modes expansions of generalized coordinates and inputs in conjunction with the variational work-energy principles that govern the physical system. The algebraic optimal control laws obtained by the Direct Method are in the form of digitally implementable generalized algebraic state feedback; to implement these control laws an algebraic state estimator must be designed. The development of such an estimator is presented by utilizing neural networks within a hybrid algebraic equations of motion-neural structure for general nonlinear systems. However, due to time constraints, computer simulations of the application of the method are presented for linear systems under deterministic, noisy and modelling uncertainty cases. As modelling uncertainty, both parameter uncertainty and model truncation have been considered. Beyond state estimation, the approach and the concepts presented in this report also form a framework for system identification via the hybrid Algebraic Equations of Motion-Neural network architecture. Extension of the method to nonlinear systems is straightforward and is left as a future endeavor.

A HYBRID ALGEBRAIC EQUATIONS OF MOTION - NEURAL ESTIMATOR
TO IMPLEMENT THE DIRECT CONTROL METHOD

Hayrani Öz

<u>Introduction</u>

The study of dynamical systems via differential equations of motion is referred to as an "indirect method." On the other hand, the study of dynamic systems without any resort to or without any knowledge of differential equations of motion is referred to as the "direct method." In the direct method, algebraic equations of motion (AEM) take the place of the traditional differential equations of motion. The AEM are obtained by using Hamilton's Law of Varying Action (HLVA) in conjunction with the assumed-time-modes expansions of the generalized coordinates (Ref. 1). The constant unknown coefficients of the assumed-basis functions in time of these expansions become the generalized (algebraic) states of the dynamic system. If there are control inputs on the dynamic system, they too can be expanded in terms of assumed-basis functions in time multiplied by constant unknown coefficients of expansion playing the role of generalized (algebraic) control inputs.

By virtue of the assumed-time-modes (ATM) expansions of the generalized coordinates and the controls, the variational work-energy quantities in HLVA can be integrated apriori in time over any time interval. This provides a set of purely algebraic equations describing the motion in terms of the constant unknown algebraic states and the algebraic control inputs for the time interval considered. The motion over arbitrarily large time intervals can be studied by simple time-marching through transition intervals invoking continuity conditions for the path of the motion.

Presently, nonlinear optimal control problems are formulated in terms of differential equations. In most cases, however, it is not possible to formulate an explicit, nonlinear optimal control law in this setting. At best, the control law must be generated numerically at the expense of much computational effort. Instead of differential equations, the method applied in this report produces an explicit nonlinear algebraic optimal control law by using algebraic equations of motion and an algebraic quadratic control performance measure. The algebraic performance measure is again obtained by representing the generalized coordinates and control inputs of a dynamic system in terms of *assumed-time-modes expansions* and introducing them into a quadratic time-integral control performance measure. The complexities associated with nonlinear systems are no longer a serious issue with this approach, and the solutions can be obtained directly in explicit form with relative ease. Furthermore, although the approach is theoretically for continuous-time systems, it readily yields digitally implementable control laws with the proper choice of ATMs for the controls.

The optimal control method via the AEM is referred to as the Direct Control Method (DCM). Implementation of the optimal direct control laws requires algebraic state information, therefore, they

must be estimated. However, the generalized algebraic states of the AEM by necessity are related to acceleration and higher-order rates of the physical coordinates which we refer to as the σ-states as well as including the velocity and displacement which are the usual x-states. In the AEM/DCM approach σ-states are the primary variables. Therefore, an estimator design appropriate for the AEM needs to be developed. In this work, we develop an estimator unique for the AEM/DCM approach by utilizing the neural-networks (NN). Once this is accomplished, the ultimate objective is to demonstrate the DCM experimentally.

Because of their algebraic structure AEM appear to be naturally compatible with the neural networks for both off-line and on-line training. Therefore, the AEM state estimator developed in this work can be viewed as a larger neural network which has a hybrid structure in the sense that part of its structure represents the available modelling information from the AEM which is fixed and unmodifiable. The other part of the network is the "innovations neural-network (INN)" which has the usual trainable NN structure. The AEM structure can be interpreted as a bias-network attached to the INN. The INN is trained to emulate or estimate the σ-states and the x-states of the physical system to be used in a DCM control law. Embedding the available modelling information within a larger NN consisting of the hybrid AEM-INN estimator is conjectured to lead to or require a trainable NN that will have an advantage over having to train a larger NN that utilizes no available modelling knowledge of the physical system.

From a broader perspective, the proposed AEM-INN architecture can also be recognized as a framework for system identification if the focus is shifted from state estimation to output estimation directly. In this regard, the combination of the AEM bias-network and the innovations neural-network, INN, constitute a "postulated identification model" of the physical system. In this form, the parameters and the size of the AEM-bias network do not have to match those of the physical system. Once (and if) the INN learns to match the output of the physical system then the assumed AEM-INN architecture of the trained system, by definition, qualifies as an identified model of the physical system. The σ-states and the x-states of the AEM-INN identification model represent realization of the internal states of the physical system.

In this report, we present the approach for general nonlinear systems within the AEM formulation and discuss several scenarios for which the AEM-INN estimator is expected to perform satisfactorily, such as the deterministic or noisy environments as well as parameter and modelling uncertainties. However, due to relatively short tenure of residence at the Laboratory, the computer simulation studies of the hybrid AEM-INN estimator presented in this report are based on linear systems. Demonstrations for nonlinear systems will be undertaken at a later time. To this end, the double-flexible link PACE (Planar Articulating Controls Experiment) structure at the Phillips Laboratory

appears to be an ideal candidate. It should be noted, however, that the full potential of the AEM/DCM approach is exposed truly when dealing with nonlinear systems.

Algebraic Equations of Motion

In the AEM setting, the generalized coordinates of a system are expanded in time, over the interval (t_0, t_1), in terms of admissible time basis functions. For n generalized coordinates, $q_r(t)$, these expansions can be written as

$$q_r(t) = \sum_{k=0}^{N_r} A_{rk}(t) a_{rk} \qquad (r=1,2,...,n) \tag{1}$$

where $A_{rk}(t)$ are admissible assumed time basis functions which are independent and continuous with continuous derivatives over (t_0, t_1) which will be referred to as the transition interval, and a_{rk} are unknown constant coefficients.

Hamilton's Law of Varying Action (HLVA) (which is not the Hamilton's Principle) can be expressed as

$$\delta \int_{t_0}^{t_1} (T+W) dt - \delta q^T \frac{\partial T}{\partial \dot{q}} \Big|_{t_0}^{t_1} = 0 \tag{2}$$

where T is the kinetic energy, W is the work, and q is the n dimensional generalized coordinate vector. Once the assumed-time-modes expansions are invoked, each term becomes a known function of time. This allows HLVA to be explicitly evaluated, thereby eliminating time, yielding a set of n algebraic equations of motion in which the coefficients a_{rk} become the unknown algebraic coordinates or states.

Similar to the generalized coordinates, the inputs, $f_r(t)$, on a system can also be expressed in terms of basis functions in time as

$$f_r(t) = \sum_{k=0}^{M_r} B_{rk}(t) \beta_{rk}, \qquad (r=1,2,...,m) \tag{3}$$

where $B_{rk}(t)$ are admissible assumed-time-modes for the inputs, and the constants β_{rk} become the unknown algebraic input coordinates or input states. When substituted into HLVA, Eq. (2), for the virtual work of the inputs, expansion (3) also allows apriori explicit integration in time. Then the resulting algebraic equations of motion are in terms of the unknown coefficients a_{rk} (algebraic states) and β_{rk} (algebraic inputs).

It proves convenient to express the HLVA in nondimensional time. For an arbitrary transition interval of motion (t_1, t_0), we define the nondimensional time τ:

$$\theta \tau = t - t_0, \quad \theta = t_1 - t_0, \quad 0 \le \tau \le 1 \tag{4}$$

In terms of nondimensional time τ the form of HLVA is:

$$\int_0^1 (\delta T + \delta W_C + \delta W_{NC} + \delta W_D) d\tau - \delta q^T \frac{\partial T}{\partial q'} \Big|_0^1 = 0 \tag{5}$$

where subscripts C, NC and D denote the work of conservative, nonconservative and damping forces, respectively. A prime ($\prime$) denotes the derivatives with respect to nondimensional time τ. Any control inputs are to be included under the nonconservative work expression. By assuming the functional forms of the energy-work expressions in terms of generalized coordinates, the general form of the AEM have been derived in Ref. (1).

The explicit mechanics of obtaining the AEM will not be presented in this report. We refer the reader to the references by Öz et al in this regard. It should, however, be noted that the whole process of obtaining the AEM for a given system can be systematized much like in the familiar finite-element method. For our purpose it will suffice to state the general form of the AEM for a general nonlinear/linear/time-invariant/time-variant system.

The general form of the AEM are (Ref. 1)

$$\left[P_L + P_N(x_0)\right] a + \left[R_L + R_N(x_0)\right] x_0 + N(a, x_0, Q_f) = 0 \tag{6}$$

where subscripts L and N denote terms that rise due to linear and nonlinear behavior in the dynamic system. The term N includes all the nonlinearities in the system in terms of the a-states; and without loss of generality any other functions such as the external forcing functions represented by Q_f. The appearance of the initial states x_0 in N are due to coupling with the a-states. Unlike their differential equations counterpart, the nonlinear-AEM have a definite, well-representable structure. This allows one to deal with and derive much more explicit expressions with the AEM. Also, it should be noted that the parametric effect of initial states x_0 in a nonlinear problem are represented explicitly by the P_N and R_N terms, a fact which is not transparent in a differential equations of motion setting. The matrices P_L and R_L are, on the other hand, due to linearities and always constant and existing parameter matrices for the given interval of time (t_0, t_1). For a time-invariant system, P_L and R_L remain constant throughout the motion regardless of the size of the transition interval. Similarly, for time-invariant systems the parametric forms of nonlinear terms P_N, R_N, N would remain invariant as well. For time-variant systems, the parameter coefficients in all of these matrices would be different from one transition interval to another. But for any one particular transition-interval (t_0, t_1) again all of these matrices would remain constant.

For control purposes, it is convenient to write the AEM (6) by separating the linear and nonlinear dynamics terms, in alternate forms.

$$Pa + Rx_0 + N = 0 \quad , \quad P = P_L + P_N, \; R = R_L + R_N \tag{7a}$$

or

$$P_L a + R_L x_0 + N = 0 \quad , \quad N = P_N a + R_N x_0 + N \tag{7b}$$

The solution procedure for the response via the AEM consists of solving the algebraic equations (6) or (7) for the unknown constants σ (σ-states) and using these constants to evaluate the assumed-time-modes expansions of the generalized coordinates, Eq. (1) for $0 \leq \tau \leq 1$ corresponding to any time t in the interval (t_0, t_1). In order to obtain a solution over an arbitrary length of time, Eqs. (6,7) can be marched in time from one transition interval to the next. This is accomplished by using the following continuity equations:

$$x_0^{(j+1)} = S_0 x_0^{(j)} + S\sigma^{(j)} \tag{8}$$

where (j) and (j + 1) represent two subsequent transition intervals. The continuity matrices S_0 and S are deduced for any specific choice of the ATM expansions by keeping in mind that the generalized velocity and displacement at the end of transition interval (j) at $\tau = 1$, are the initial velocity and displacement for the next transition interval (j + 1) at $\tau = 0$. For example, for simple power series expansions in time of the generalized coordinates q, the continuity matrices, S_0 and S, are (Ref. 1)

$$S_0 = Block \ Diag \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad , \quad S = Block \ Diag \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 2 & 3 & \cdots & N_r \end{bmatrix}$$

The control inputs on a dynamic system are external loads designed to satisfy certain performance criteria. Similar to the generalized coordinates, the control input functions can also be represented as expansions in assumed-time-modes. Denoting the control forces by $f^c(t)$, for m-inputs, the control input expansions can be written as

$$f_r^c(t) = \sum_{k=0}^{M_r} B_{rk}(t)\beta_{rk} = B_r(t)\beta_r \qquad (r=1,2,\ldots,m) \tag{9}$$

where $B_{rk}(t)$ are admissible assumed-time-modes for the control inputs and β_{rk} are constant unknown algebraic control input coordinates that are determined to meet the control objectives. In Eq. (9), $B_r(t)$ and β_r are M_r-component vectors of ATM and algebraic input coordinates, respectively with obvious definitions. The control inputs can be added to the AEM simply by considering their virtual work and transforming to the domain of σ-states. With the additional control term, the AEM take the form:

$$P\sigma + Rx_0 + N + Q\beta = 0 \tag{10}$$

where Q is the algebraic control influence matrix from the algebraic (generalized) control input vector β to the σ-states.

In general, the assumed time modes B_{rk} for the inputs can be taken to be the same as the assumed-time-modes A_{ij} for the generalized coordinate(s). However, another simple choice is one in which $M_r = 1$, $k = 1$ and $B_{rk}(t) = 1$ $(r = 1,2,\ldots,m)$. That is, each input is a single term expansion with constant assumed-time-mode shape over each interval $0 \leq \tau \leq 1$. This is indeed very well suitable for (digital) implementation of the control inputs where each control coordinate β_{rk} corresponds to the value of the physical control input. This particular ATM for controls literally corresponds to a zero-

order-hold in sampled-data systems. Indeed, things do not need to be any more complicated. Simplicity is sufficient and leads to an elegant implementable form for control inputs with the Direct Control Methodology.

It now remains to obtain the algebraic control inputs β to satisfy desired controller performance criteria. Possibly one would be interested in finding the control inputs β as functions of α, the algebraic states, which would by definition represent state feedback control. An explicit nonlinear feedback control solution for the nonlinear dynamics represented by the AEM, Eq. (10) is given in Ref. (2) by using the quadratic regulator performance measure typically used in optimal control theory. For completeness and instructive purposes we will briefly give the steps that lead to an explicit optimal solution. The direct optimal control methodology described in this report is applicable to multi-degree-of-freedom, multi-input systems in which q represents an n-component vector of generalized coordinates.

Direct Optimal Nonlinear Control

Since the structure of the AEM are well-defined for the most general dynamic systems, the form of the control laws presented in this section is also the most general and they also reduce to linear laws readily if the system is linear. Hence, linear controllers are inclusive within the nonlinear control presentation.

To design a direct optimal control law for the system we choose a control performance measure. As an example of wide interest and applicability, the usual quadratic regulator form is selected as a suitable control design performance index (CDPI):

$$J_t = \frac{1}{2} \int_{t_0}^{t_1} \left[q^T W_q q + \dot{q}^T W_{\dot{q}} \dot{q} + f^{cT}{}_{(t)} R_f f^{cT}{}_{(t)} \right] dt \tag{11}$$

where W_q, $W_{\dot{q}}$, are positive semi-definite state weightings and R_f is the positive definite control weighting matrices. For simplicity, cross products between the displacements and velocities are not considered. The CDPI can be transformed to the domain of the AEM by introducing the ATM expansions (1) and (9) for the generalized coordinates and the input functions. In this transformation, since the time dependence of the terms are known by virtue of the ATM, the time integrals in the CDPI can be evaluated apriori to yield the equivalent algebraic performance criterion $J_t(q, \dot{q}, f^c) = J(\alpha, \beta)$:

$$J(\alpha, \beta) = \frac{\theta}{2} \left[\alpha^T (D + V) \alpha + 2\alpha^T (D_0 + V_0) x_0 + \beta^T F \beta \right] + O(x_0) \tag{12}$$

where the matrices D and V are apriori evaluated matrices related to displacement and velocity weightings integrated with the ATM expansion shape matrices for the coordinates. Similarly, F is also an apriori evaluated matrix related to control weighting matrix integrated with the ATM expansion shape matrix for the control inputs. $O(x_0)$ denotes terms due to x_0 only and they do not affect the optimization of $J(\alpha, \beta)$.

The optimal control problem now becomes a completely algebraic optimization problem. $J(\alpha,\beta)$ can be optimized subject to the AEM as constraint equations. For fixed final time and free final states, the explicit nonlinear optimal control law β^* for the transition interval (t_0, t_1) is (Ref. 2)

$$\beta^* = F^{-1}Q^T\left[Z_N^T + P^T\right]^{-1}\left[(D+V)\alpha + (D_0+V_0)x_0\right] \quad , \quad Z_N = \left[\frac{\partial N_i}{\partial \alpha_i}\right] \tag{13}$$

is the Jacobian due to nonlinearities N. The optimal nonlinear control law has the form

$$\beta^* = G(\alpha,x_0)\alpha + G_0(\alpha,x_0)x_0 \tag{14a}$$

where

$$G(\alpha,x_0) = F^{-1}Q^T\left[Z_N^T(\alpha,x_0) + P^T(x_0)\right]^{-1}[D+V]$$

$$G_0(\alpha,x_0) = F^{-1}Q^T\left[Z_N^T(\alpha,x_0) + P^T(x_0)\right]^{-1}[D_0+V_0] \tag{14b,c}$$

are the control gain matrices. The α and x_0 dependencies of the control gains are solely due to nonlinearities characterized through the Jacobian and the contributions to P due to nonlinearity in the system as defined by Eqs. (6,7). All other matrices involved in the control gain matrices are constant matrices computed apriori by performing the time-integrals thanks to the ATM expansions. The optimal control law Eq. (14a) represents an explicit solution to a nonlinear optimal control problem obtained by simple algebra. The solution given here is in direct contrast to traditional differential equations of motion approach which can never yield explicit solutions for optimal control but can only offer a formalism for numerical solution techniques involving two point boundary value problems and intermediate differential equations.

The explicit direct optimal control law is quite general and readily reduces to an optimal solution for a linear system. Indeed, if we use the alternate form of AEM, Eq. (7b), the optimal control law can be written as

where

$$\beta^* = F^{-1}Q^T\left[Z_N^T + P_L^T\right]^{-1}\left[(D+V)\alpha + (D_0+V_0)x_0\right] \tag{15a}$$

$$Z_N = \left[\frac{\partial N_i}{\partial \alpha_j}\right] \quad , \quad N = N + P_N\alpha + R_N x_0 \tag{15b}$$

is the Jacobian that includes all the nonlinearities. When there are no terms due to nonlinearities in the system the Jacobian vanishes and the optimal control law reduces to a linear control law by default. This is a necessary requirement for a nonlinear control law.

Finally, replacing β in Eq. (10) with β^*, Eq. (14a or 15a), produces a closed-loop-control form of the AEM, which can be rearranged in terms of the unknown optimal algebraic states, α:

$$\alpha = -P^{-1}Rx_0 - P^{-1}N - P^{-1}Q\beta^* \tag{16}$$

This nonlinear equation represents the solution form of the closed-loop controlled nonlinear system and is used to simulate the closed-loop controlled response. For nonlinear systems, because Eq. (16) is

nonlinear in σ, an iteration must be used, for simulation purposes, to solve for σ. Note that, this is not an iteration to obtain the control law. The control law is already given in explicit form by Eq. (14).

On the other hand, for linear systems, the closed-loop system response solution is immediately available without any need for simulation iteration:

$$\sigma^* = -(P + QG)^{-1}(R + QG_0)x_0 \quad ; \quad \beta^* = \left[G\sigma + G_0 x_0\right] \tag{17a,b}$$

$$\sigma^* = -P_{CL}^{-1} R_{CL} x_0 \tag{17c}$$

where definitions of P_{CL} and R_{CL} are evident as the closed-loop parameter matrices.

We now turn to implementation of the DCM.

A Hybrid AEM and Neural Estimator

Since the optimal control yields full-state feedback control laws in both σ and x states, to implement it, one must have knowledge of the system states required in the control law. In general, the σ and x-states of the AEM may correspond to more abstract coordinates, such as modal coordinates. Hence, one needs to estimate them by using outputs/measurements from the system that correspond to physical quantities. To this end, we propose to design an estimator for the AEM in conjunction with a neural network block employed by the estimator. In this set-up, the function of the neural network (NN) is envisioned as being an "innovator" that generates output in the estimator to correct the error between the estimated states and the actual states. This "innovations neural network (INN)" will be driven by the measurements and estimated outputs at the current time and possibly at a specified number of previous times (see Fig. 1). If an analogy is appropriate, the INN will play the role of an estimator gain block, that will train itself, that is determine its weights, to minimize the state estimation error.

The motivation for designing a neural-based estimator for the AEM is based on the following observation. Consider, for example, the linear AEM. For given initial states and forcing functions, as inputs, the σ-states represent outputs obtained by a simple matrix multiplication the elements of which are the system parameters, playing the role of a weighting matrix. Thus, this algebraic system is equivalent to a linear NN architecture with unit operator activation functions and zero biases. For nonlinear systems, one can interpret the nonlinear terms as being nonlinear σ-state feedback terms introduced (summed) to the outputs of the linear dynamics which were already observed to be in a simple NN form. In addition, in the nonlinear case, there are feedforward connections from x_0 due to the $R_N(x_0)$, $P_N(x_0)$ terms. Thus, the AEM naturally appear to have a structure that is compatible with neural networks. The combination of the AEM with a "NN innovations" block also is expected to benefit the neural-network as will be discussed in the sequel. Hence, a symbiotic relationship is to be established between the AEM and the NN.

We postulate the following form for an AEM estimator dynamics which also includes an INN dynamics:

$$P\hat{a} + \hat{R}x_0 + N(\hat{a}, \hat{x}_0) + Q_f - I_{NN} = 0 \qquad (18a)$$

and for completeness note the AEM for the system:

$$Pa + Rx_0 + N(a, x_0) + Q_f + w = 0 \qquad (18b)$$

in which (^) denotes the estimated states and a forcing function Q_f has been included explicitly. The term I_{NN} in the AEM estimator is the output of the INN which is to be designed to reduce the state errors between the estimator and the dynamic system. In the above, for generality a plant noise vector w has been included in the AEM, thus presenting a more realistic situation.

On the other hand, as a different scenario, it must be noted that the estimator matrices P, R and N in (18a) can be different from those of the system in (18b) both in terms of parameters and size. In this case, when (if) the I_{NN} network is (can be) trained to map the estimated output of (18a) to the output of the system (18b) instead of matching the a and x-states, then the estimator dynamics of Eq. (18a) becomes an "AEM-Neural identification model" for the physical system. In this case the $\hat{a}$ and $\hat{x}$ states of Eq. (18a) rather than being the estimated states become the surrogate or equivalent realization states of the physical system (18b). Hence, with proper interpretation, Eq. (18a) can be viewed either as a state-estimation model or an assumed identification model. In this context, the reader is cautioned that the use of the word "estimator" in this report can imply either state-estimation or model identification.

The output equations for the AEM can be considered both for the beginning or ending of the j-th transition interval, in the form

$$y^{j+1} = Cx_0^{j+1} = C\left(S_0 x_0^j + Sa^j\right)$$

where the superscript (j + 1) indicates the beginning of the (j + 1)-th transition period or the end of the j-th transition period. Thus, the far right side of the output equation involves the a-states explicitly by introducing the continuity equation (8). We shall, for simplicity of notation, prefer the output equation in terms of x-states only. Further, assuming that the Eqs. (18) pertain to the j-th transition interval, for estimator design we shall use the outputs at the beginning of the j-th interval. Thus, we consider, the output equation as

$$y^j = Cx_0^j \qquad \text{or} \qquad y = Cx_0$$

the interval designation will be dropped unless a confusion may arise. For generality, we shall assume that the measurements are also contaminated by a noise vector v to write

$$y = Cx_0 + v \qquad (18c)$$

The estimate of the measurements will be in the form

$$\hat{y}^j = C\hat{x}_0^j = C\left(S_0\hat{x}_0^{j-1} + S\hat{\sigma}^{j-1}\right)$$

The NN innovations output I_{NN} will use the current and a finite number of past measurements and their estimates as the input to the NN:

$$I_{NN} = I_{NN}(y, \hat{y}) \tag{19a}$$

Note that, by design, no system input Q_f will be provided to the INN. The particular architecture of the estimator and the NN training flow diagram is shown in Fig. 1. As discussed in Ref. (4), within the context of observability of the AEM observation of the x-states is essential and the observability of the σ-states follow automatically from the AEM. Hence, although we wish to estimate both x_0 and σ simultaneously, we base the error performance measure of the NN design, J_{NN}, on the state estimation error in x_0 and write

$$2J_{NN}^{j+1} = (x_0 - \hat{x}_0)^T Q_N (x - \hat{x}_0) \qquad Q_N^T = Q_N > 0 \tag{19b}$$

where Q is an estimation error performance weighting matrix for the NN. When $\hat{x}_0$ converges to x_0; convergence of $\hat{\sigma}$ to σ is guaranteed. Similar to Eq. (19b) an σ estimation error performance can also be stated separately. Alternately, the NN performance measure can be based on the output estimation error directly:

$$2J_{NN}^{j+1} = (y - \hat{y})^T Q_N (y - \hat{y}) \qquad Q_N^T = Q_N > 0 \tag{19c}$$

In both (19 b,c), the outputs and states are considered to be at the beginning of the (j + 1)th transition interval (or the end of the j-th transition), indicated by the superscript (j + 1). Hence, there is a one transition interval delay between the J_{NN} and I_{NN} terms. Therefore, the INN has a recurrent structure.

All of the terms in the AEM estimator dynamics described by Eq. (18a) can be regarded as a single larger neural network as shown in Fig. 1 in which the first four-terms represent an unmodifiable bias network structure and the last term I_{NN} represents the part of the network that is modifiable and therefore trainable. The unmodifiable part consisting of AEM structure constitutes known or assumed modelling information about the system that is going to be estimated or identified; hence, this information is embedded into the overall NN structures as the unmodifiable part. The remaining part of the NN then is trained to learn to eliminate the state estimation error or the output estimation error only by using the measurements. It is therefore, expected that the NN will not have to be burdened both in training and in size as much as if one were to completely ignore the available model information about the system and attempt to design a fully trainable NN utilizing no apriori information about the dynamic system. Whereas the AEM utilize the NN to obtain the σ and x-state estimates to employ the control laws; in turn, due to their algebraic nature they become a compatible part of a larger NN estimator in which only the innovations block has to be trained as a smaller, perhaps faster and more able neural partner.

From a broader perspective, the function of the NN innovations in (18a) is to learn any differences between the dynamics of (Eqs. 18a,b) and compensate for it in terms of minimizing an output error measure between the two systems. In that, any NN itself having unique structural (architectural) attributes, is theoretically indifferent to the specific form or the nature of the terms that describe the internal dynamics of the systems that it is attached to learn. Given the proper inputs, the focus of a NN is strictly on matching its output to that of the dynamic system in terms of the parameters of its unique structure. With this caveat, we make the following observations for Eqs. (18a,b).

First disregarding the noise on the plant and the measurements, if the parameter matrices P, R, Q_f and the forms of the nonlinearities N are the same in both (18a and 18b) then, inherently the INN will only learn to eliminate the measurements error. When this is accomplished it should produce no output since there will be no need for corrections. When there are noises w and v on the plant and the measurements, the INN will still train as best as it can since it is focused only on the output error and does not require explicit information about what constitutes the system output, noisy or not. It may be surmised that, it will try to learn the noise as well if the noise appears to have a structure.

When there are differences in the structures and/or parameters of the dynamic-systems represented in Eq. (18a,b), the INN will again try to minimize the output error between the two systems which is the only task it is charged with and the motivation for its learning. Consequently, it will also try to learn to make up for the modelling differences between (18a and b). To this end, it will try to map the structural differences between the two models in its own terms. In this case, the INN will have to provide continuous correction for the differences between the estimator and the dynamic system. Once the INN converges, Eq. (18a) becomes an "identified model" of the physical system.

The validity of the above conjectures will be illustrated by examples in the next section. In summary, the form of the postulated hybrid estimator dynamics Eq. (18a) together with the INN and the formalism of the AEM is anything but restrictive. Furthermore, when interpreted properly with the properties of neural-networks one indeed obtains very broad and powerful ranges of possibilities of practical interest to study including both state estimation and system identification.

In reality, although one may not have perfect model information about physical systems; one is never clueless about appropriate physically reasonable models of the systems either. In this case, it is reasonable to exploit whatever information one has about the system and utilize it together with an INN as in Eq. (18a) to leave the discrepancies for the INN. In our case, the particular model structure we choose to work with the NN is the AEM given by Eq. (18a). Note that Eq. (18b) serves as a convenient model of the physical system to be used as a NN trainer, therefore, it does not have to be in the form of AEM. However, in Eq. (18a) we precisely choose to use the AEM form for the potential the AEM formulation offers for the control of general dynamic systems via the DCM as discussed previously. From the perspective of system identification, the unmodifiable AEM part of the estimator

(18a) can be based on the best available model of the system or on a model related to the system to be identified; all parameter modifications take place only within the INN structure. Once the postulated AEM-neural estimator of Eq. (18a) is trained to emulate the physical system then it can be used faithfully for all design purposes. To this end, since the nonlinear structure of the trained INN and its parameters are known, from the perspective of AEM/DCM, it effectively becomes part of the nonlinear terms. That is, with a control input, and the INN in the loop, by redefining the nonlinear term and showing its arguments explicitly, Eq. (18a) can be rewritten as

$$P\dot{a} + R\dot{x}_0 + N\big(N(\dot{a}, \dot{x}_0), Q_f, I_{NN}(y, \hat{y})\big) + Q\beta = 0 \tag{20}$$

which is exactly in the form of Eq. (10) where $\dot{a}$, $\dot{x}_0$ are now the same as a, x_0 since Eq. (20) emulates the physical system. Design of optimal controls for the nonlinear estimated/identified system of Eq. (20) by using the DCM will not pose a problem now.

Application of the Method

The hybrid AEM-Neural estimator configuration originated because of the author's plan to implement the DCM experimentally on a vibration control problem of interest to the Air Force at the Phillips Laboratory in the near future. One of the experimental models in the laboratory is a linear dynamic flexible beam with a number of surface mounted piezoelectric actuators and sensors. Another experimental model is the nonlinear double-flexible-link PACE (Planar Articulating Controls Experiment) structure. As an immediate application of the concepts presented in the previous sections, we focused on a linear system and left the demonstrations for nonlinear systems as a future undertaking. However, it should be realized that nonlinear systems pose no difficulty for the AEM/DCM (in fact, the approach thrives for nonlinear systems) and certainly neural-networks are also most capable of dealing with nonlinear systems. Therefore, the combination of the AEM/DCM with neural networks for nonlinear systems appears to offer attractive possibilities as can be inferred from the discussions in the preceding sections. Our consideration of the linear systems at this point is a matter of convenience. Furthermore, a point of immediate interest is to verify the observations regarding the possible capabilities of the hybrid form of estimator dynamics of Eq. (18a) elaborated in the previous section without having to get involved in the excessive computer code development during the relatively short period of residence at the Laboratory.

The neural-network innovations block of Fig. 1 and Eq. (18a) is designed by using two hidden layers with the back propagation algorithm. For the first hidden layer, hyperbolic tangents were used as activation functions. Since the amplitudes of these functions are limited to (-1,1), the second hidden layer used unit activation functions to be able to accommodate the unknown range of signals that will be propagating through the network. The weight adjustment procedure for the network by

back propagation is standard and will not be repeated here. Referring to Fig. 1, it is noted that the INN performance error is not immediately at the INN output layer and is ahead of it through the P^{-1} and the continuity blocks. For back propagation, the error to be backpropagated to the INN output layer can be shown to be

$$\delta_k^{out} = \sum_{i=1}^{2n} \frac{\partial J_{NN}}{\partial \hat{x}_{oi}} \frac{\partial \hat{x}_{oi}}{\partial I_k} \frac{\partial I_k}{\partial NET} \quad k=1,\ldots,n_\sigma \tag{21a}$$

where n and n_σ are the number of generalized coordinates and the total σ-states and I_k denotes the k-th output of the INN and J_{NN} is given by Eq. (19b). Carrying out the operations one obtains

$$\delta_k^{out} = -\sum_{i=1}^{2n} Q_{ii} \epsilon_{oi}(SP^{-1})_{ik} f_k' \; ; \tag{21b}$$

$$\epsilon_{oi} = x_{oi} - \hat{x}_{oi}, \quad f_k' = 1 - Tanh^2(\cdot)$$

in which the error weighting matrix Q_N is assumed diagonal. The back propagation algorithm proceeds as usual by using the above form of δ_k through the output layer of the INN. Similarly, instead of state estimation error, if the measurement estimation error Eq. (19c) is used for the INN performance, the back propagation error term will be

$$\delta_k^{out} = -\sum_{i=1}^{m} Q_{ii}(y - \hat{y})_i \, (CSP^{-1})_{ik} f_k' \tag{21c}$$

Thus, in both cases, one simply pushes backwards the performance errors through the fixed-forward structure of the estimator to the innovations INN block and the back propagation continues there on. Both of the algorithms in Eq. (21) can be used for on-line training as depicted by Fig. 1.

The following computer simulations have been performed as proof-of-concept studies.

A Single-Degree-of-Freedom System: A simple spring-mass system with the mass and stiffness parameters of 1 and 10, respectively, was considered. Initial displacement and velocity were taken to be 1 and 0. A single displacement measure was taken. The initial state estimates $\hat{x}_0$ were taken to be zero. No initial states for the σ-states are needed. Two σ-states corresponding to acceleration and rate of acceleration (jerk) as coefficients of parabolic and cubic (in time) assumed-time modes expansion in Eq. (1) were sufficient to simulate the system faithfully for a transition interval of (t_0, t_1), 0.01 seconds. The INN used 3 inputs corresponding to the current and past two time-delayed measurement errors. The two hidden-layers had each 20 neurons and the output-layer had as many outputs as the total number of σ-states to be estimated.

The learning rate was $\eta = 0.25$ with a momentum gain of 0.5. In each case studied, it took about 3,000 transition intervals for the NN to learn the states. Figures 2a show the displacement and its estimate in the closed-loop system when a linear optimal direct control law is implemented by using the estimates $\hat{b}$ and $\hat{x}$. The weightings for the quadratic regulator design were $w_q = 1$, $w_{\dot{q}} = 50$,

$R_f = 0.1$. In these deterministic cases, after learning successfully, the innovations NN ceased to yield output as expected since it was using measurement errors as inputs and perfect knowledge of the plant model was assumed.

Next, the estimator was trained in the presence of both zero-mean Gaussian plant and measurement noises of unit variances with gains of 10 and 0.2, respectively. Figures 2b show the closed-loop responses for the displacement and its estimate when the noise trained estimator implemented the optimal direct control law that was designed for the deterministic case. It should be noted that the actual system response that is realized is the smoother "displacement" curve on the left although the noisy a and x-estimates like the "displacement estimate" were fedback to implement the control law.

The third case considered parameter uncertainty, hence the estimator was essentially an identifier. It was assumed for the physical system of Eq. (18b) that the spring stiffness was 15 whereas in Eq. (18a) for the estimator an incorrect model with the (previous) stiffness of 10 was retained, however, both systems had the same mass. Again the estimator trained successfully in spite of the parameter error. As surmised, the innovation NN, however, provided continuous output to compensate for the mismatch between the model and the assumed model used in the estimator design. It can be said that the innovations NN emulated this mismatch in its own structure. Figures 2c show the closed-loop responses for the displacement and estimates of the actual system controlled by the AEM-INN estimator which was based on the wrong (or a different) model of the system.

A Multi-Degree-of-Freedom System: A 3-mass ($m = 2,4,1$), and 3-spring ($k = 10,10,10$) system restrained at one end was considered. A two-term ATM expansion for each coordinate with a transition interval of 0.01 secs. was used as the AEM model. The neural network used the current and past four measurement sets as inputs each consisting of three displacement measurements. The two hidden-layers had 35 and 25 neurons with a learning rate of 0.2 and the momentum gain of 0.5. It took about 30,000 transition intervals for the AEM-INN to estimate the states. Figure 3a shows the closed-loop responses of the coordinates via an optimal DCM control law using three control inputs with weights $w_q = [1]$, $w_q = [50 \ 50 \ 1]$, $R_f = 0.1$, implemented by the AEM-INN estimator. Note that although the estimates appear oscillatory, the actual system response realized is smooth as shown by the figure on the left.

Next, we considered the 3rd order modal dynamics of the system. However, our interest was in a reduced-order model of the system for estimation and control purposes. We assumed the reduced-order model to be the lowest mode and trained the INN to estimate the states of the first mode. The INN used as input the measurement error between the output of the full-order (3rd order) system and its estimated output based on the reduced-order (1-mode) model. The same neural network parameters given above were used. The estimator learned the first mode in about 2,000 transition intervals. In

contrast, previously, for the training for the full-order model, it was the first mode that was hardest to learn which had delayed the learning to 30,000 steps. Figure 3b shows the output due to the first mode and its estimate and the actual plant output of the 3rd-order system that was fedback to the INN. Effectively, the NN was able to learn to filter out that part of the response we are not interested in (the higher two modes) and give us the desired estimate for the reduced-order model. We have thus constructed a modal neural estimator targeted for a specific mode. The objective of this example was to demonstrate that the AEM-Neural estimator can be trained properly even in the case of model truncation which is another source of uncertainty besides parameter errors. Certainly, the reduced-order model can be any subset of the physical system. Note that since we demonstrated the case also for parameter uncertainty, one may presume that the parameters in the reduced-order system do not have to be known exactly either.

Concluding Remarks

In this work, we showed the viability of designing a hybrid algebraic equations of motion-neural estimator to implement an algebraic optimal control approach referred to as the Direct Control Method (DCM). The novel feature of the algebraic equations of motion (AEM) based DCM is that no resort to any differential equations is made nor any knowledge of differential equations is required. The AEM/DCM presents a unified approach to dynamics and control of general nonlinear, linear, time-variant, time-invariant systems. The AEM/DCM approach is naturally compatible with neural-network architectures and the advantages of both theories can be combined to develop a unique and powerful method for dynamics and control of complex dynamic systems with on-line capabilities. Therefore, the implementation of the DCM is studied by using a neural-network based estimator. All simulations illustrated in this report operated in on-line mode.

The hybrid neural network estimator uses available modelling information (which does not have to be exact) about the system as an unmodifiable part and the remaining part of its structure represents a trainable innovations neural-network (INN) which maps the differences between the system and the estimator. In this regard, the proposed estimator structure can also serve as a system identification model. Although the concepts are proposed for general nonlinear systems, during the relatively short tenure of the author's residence at the Laboratory, it was only feasible to present, in this report some proof-of-concept computer simulations for single and multi-degree-of-freedom linear systems. We demonstrated that the hybrid AEM-Neural estimator can learn to estimate or emulate the generalized algebraic states of the system in both deterministic and noisy environments and with modelling errors as parameter uncertainty or model truncation in the estimator. Extension of the method to nonlinear systems is straightforward and is left as a future endeavor, to exploit the full potential of the DCM and the neural-networks, and the theoretical and practical implications of the

25-17

ideas illustrated herein. To this end, the double-flexible-link nonlinear PACE (Planar Articulating Controls Experiment) structure at the Phillips Laboratory appears to be an ideal candidate.

References

1. Öz, H. and Adigüzel, E., "Hamilton's Law of Varying Action, Part I: Assumed-Time-Modes Method," *Journal of Sound and Vibration* (to appear, Sept. 1994). Also, AIAA Paper 90-1169, "General Formulation of Algebraic Equations of Motion-Assumed Time Modes Method," Procs. 31st SDM Conf. April 1990, Long Beach, CA, pp. 2381-2389.

2. Adigüzel, E. and Öz, H., "Direct Optimal Control of Nonlinear Systems via Hamilton's Law of Varying Action," AIAA Guidance, Navigation and Control Conference, New Orleans, LA, 1991, Paper No. 91-2808. Also, *Journal of Dynamic Systems, Measurement and Control* (to appear).

3. Öz, H. and Adigüzel, E., "Direct Optimal Control of Linear Systems via Hamilton's Law of Varying Action," AIAA Guidance, Navigation and Control Conference, New Orleans, LA, 1991, Paper No. 91-2647. Also under the title: "Hamilton's Law of Varying Action, Part II: Direct Optimal Control of Linear Systems," *Journal of Sound and Vibration* (to appear Sept. 1994).

4. Adigüzel, E. and Öz, H., "Controllability and Observability of Algebraic Equations of Motion by Hamilton's Law of Varying Action," AIAA Paper No.: 94-3584, Procs. of Guidance, Navigation and Control Conf., Scottsdale, AZ, August 1994.

5. Wasserman, P.D., *Advanced Methods in Neural Computing*, Van Nostrand Reinhold, N.Y., 1993.
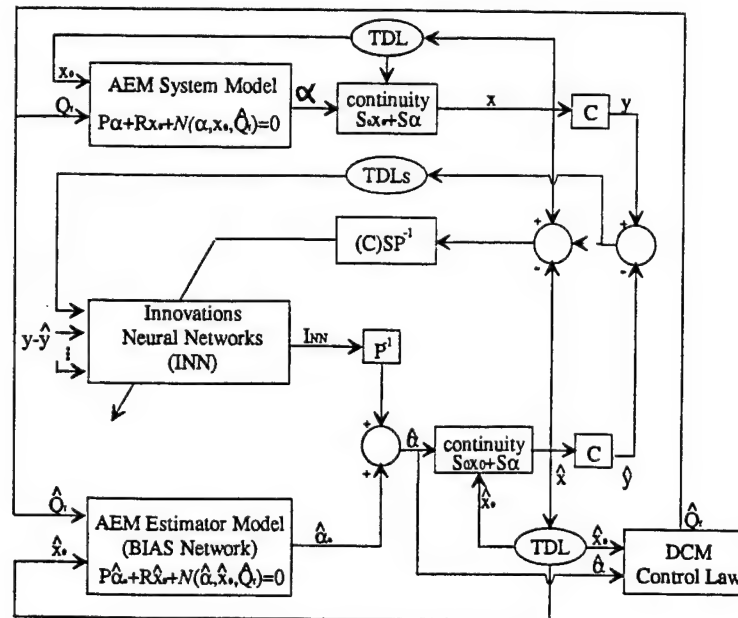
Figure 1:   Hybrid AEM-INN Estimator Functional Flow Diagram
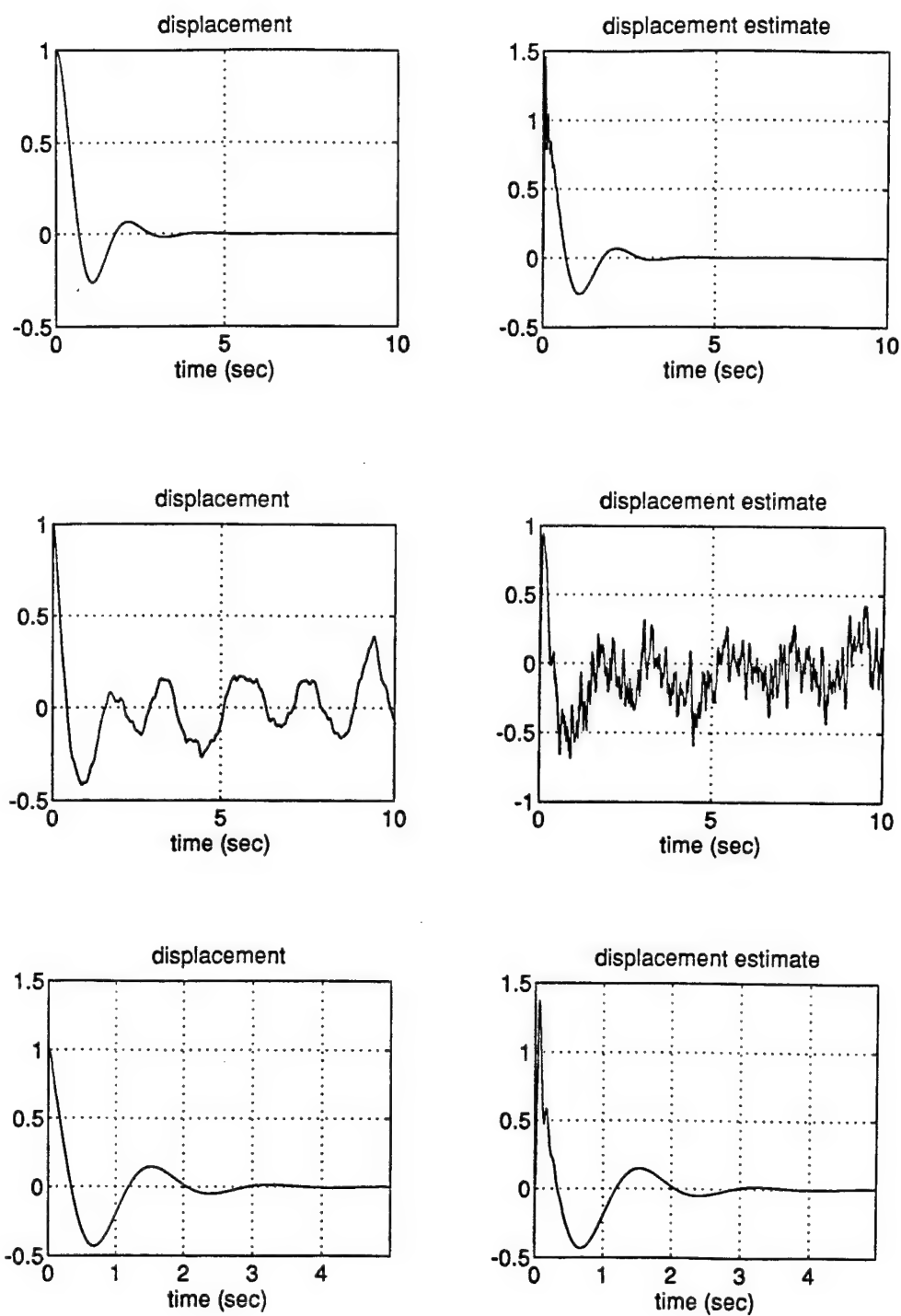
Figure 2: Single-Degree of Freedom System Controlled by the AEM-INN Estimator

a) Deterministic (top) b)Plant and Measurement Noise (middle)

c) Parameter Uncertainty (bottom)

25-19
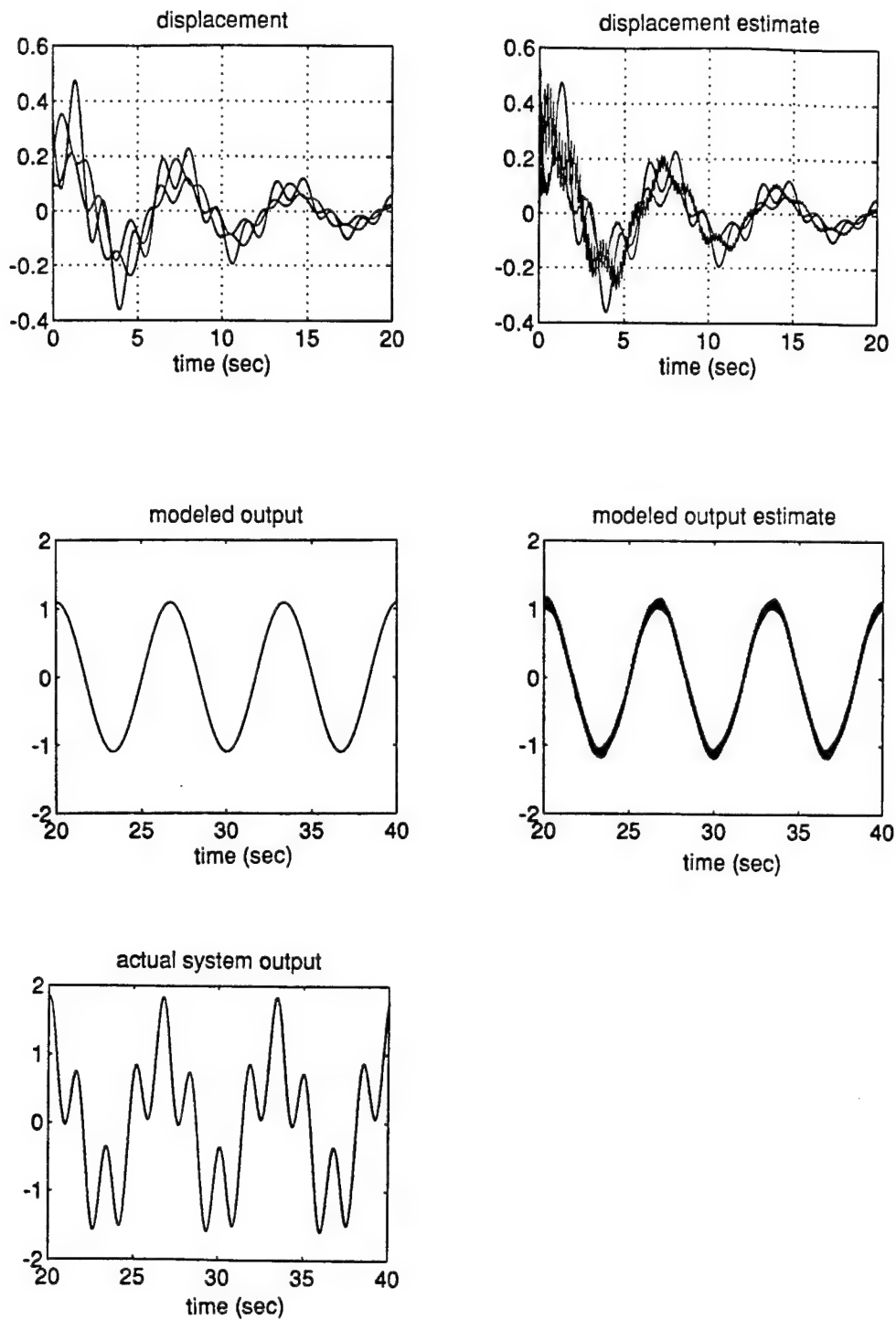
Figure 3:   3-Degree-of-Freedom System a) Controlled by the AEM-INN Estimator (top)
             b) Selective Mode-1 Estimation with Modal Truncation (middle, bottom)

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER:
THEORY AND IMPLICATIONS FOR GROUND-BASED INTERFEROMETRY

Sudhakar Prasad
Associate Professor
Center for Advanced Studies and Department of Physics and Astronomy

University of New Mexico
800 Yale Blvd. NE
Albuquerque, New Mexico 87131

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER: THEORY AND IMPLICATIONS FOR GROUND-BASED INTERFEROMETRY

Sudhakar Prasad
Associate Professor
Center for Advanced Studies and Department of Physics and Astronomy
University of New Mexico

Abstract

We developed a variational, scalar-field treatment of the guided modes and their propagation constants for a multiple-core optical fiber consisting of many single-mode single fiber cores closely spaced inside a common cladding material. The variational procedure, when supplemented with symmetry considerations, produced without great effort the modes and their wavevectors for a highly symmetrical arrangement of an arbitrary number of cores. By employing these modes, we computed an expression for the coupling efficiency of an arbitrary wavefront focused into such a fiber. We argue that such a fiber should be more strongly coupled than the standard single-core single-mode fiber to an arbitrary wavefront due to a piecewise coupling of the wavefront to the multiple-core fiber. The improved coupling efficiency is illustrated for a variety of turbulence-induced wavefront distortions typically seen by ground-based imaging systems. Coupling efficiency improvements by about an order of magnitude are possible under severe turbulence conditions, when compared with the performance of standard single-core single-mode fibers. Also, by making use of evanescent-field interaction among cores which causes a nontrivial evolution of core excitations and mutual coherences with propagation, one can generate sub-beams of light, one for each core, with comarable excitations and high mutual coherence. Such sub-beams from different telescopes of an interferometric array, when combined with each other, will generate fringes of high quality with rather modest co-phasing and adaptive corrections. Furthermore, the MC fiber design is rather insensitive to the conditions of turbulence, which makes MC-fiber couplers very attractive for ground-based interferometry.

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

I. Introduction

The principal advantages of using single-mode (SM) optical fibers in interferometry derive from the fact that such fibers serve as spatial filters[1] that reject any spatial noise imposed on wavefronts. The perfect spatial coherence of the guided mode is manifested in SM-fiber interference experiments in high fringe visibility, which is independent of imposed noise or seeing conditions,[2] rendering SM fibers as essentially ideal optical couplers for interferometry. The only problem is that the sensitivity with which fringes can be detected can be quite low due to a highly inefficient coupling process. To remedy this problem, we have proposed the use of a different kind of fiber, one with multiple cores, which is expected to be better matched to an arbitrary wavefront than the standard single-core fiber. By employing a multiple-core (MC) fiber, one can retain all of the advantages of a SM fiber that derive from perfect spatial coherence of a single guided mode, but at the same time can significantly boost the coupled power fraction. The details of this work are described completely in Refs. 3 and 4, which will soon be submitted for publication. Here we shall only present the most salient aspects of this work.

The chief objective of this paper is to demonstrate unequivocally vast improvements in coupling efficiencies possible with a MC fiber. We shall briefly review the general variational analysis we developed over the summer to address the guided modes of a MC fiber, on which the calculation of coupling efficiencies is predicated. The particular MC fiber we have studied in detail has 7 cores arranged in a highly symmetrical hexagonal geometry, in which one core is surrounded by 6 other, regularly spaced cores, as shown in Fig. 1. All of the nearest-neighbor separations are identical for this particular geometry. This case is of special relevance to the proposed 7-element CHARA interferometer.[5] One can envision an alternative design of the CHARA array in which each telescope beam is divided into 7 sub-beams for 21 distinct pairwise beam combinations (42 sub-beams), with the remaining 7 sub-beams (1 sub-beam per telescope) for fringe tracking. By coupling light in the focal plane of each telescope to a 7-core fiber, one obtains simultaneous excitation of 7 individual cores with an inner-outer core excitation ratio that can be altered, via evanescent fields, by allowing light to merely propagate in the MC fiber. After a certain propagation distance when requisite relative excitation has taken place, the cores of the fiber can be spatially separated into 7 individual single-core single-mode (SCSM) fibers, thus producing 7 sub-beams as desired.

A highly desirable feature of our MC fiber is the insensitivity of its optimal design parameters relative to changing conditions of turbulence. In other words, a MC fiber once designed and fabricated is expected to remain optimized for large variations in seeing and other noise conditions. We shall see this in detail in all our results on coupling and propagation. For example, the core radius for optimal coupling will turn out to be roughly half the Airy-disk radius corresponding to diffraction-limited imaging, *independent* of the size of the pupil relative to the Fried coherence diameter r_0 characterizing atmospheric turbulence.[6] As we have indicated in Ref. 3, the fundamental reason for the robustness of the optimal design lies in the nature of the coupling of fields into the fiber, which depends only on the correlation properties of the focused field, not on its seeing-dependent intensity profile. Not surprisingly, this robustness is

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

already present in the much discussed coupling into a standard SCSM fiber. We shall also see that the propagation distance required for achieving a certain degree of relative core excitation or coherence is also independent of the conditions of atmospheric turbulence, although for a different reason. As we shall examine in detail later, this independence results from the fact that the statistics of turbulence induced distortions of the wavefront are spatially isotropic and homogeneous, which implies certain mode-mode correlations in the guided field.

II. Review of the Variational Method

The variational method that we have adopted to address the modes and their propagation constants for a MC fiber is described in detail in Ref. 3. Here we only present a summary of the method. In the weak-guiding regime, in which the core and cladding refractive indices, n_{co} and n_{cl} differ very little: $\Delta n^2 \equiv n_{co}^2 - n_{cl}^2 << n_{co}^2$, the modes of the MC fiber at frequency ω may be taken to be linearly polarized with an amplitude distribution $U_\lambda(\vec{\rho})$ (λ: mode index) that can be expressed accurately as a linear combination of identical excitation profiles located about the individual core centers $\vec{\rho}_j$. A suitable core excitation profile is the Gaussian form, $u(\rho) = \exp(-\rho^2/2\rho_0^2)$, so that $U_\lambda(\vec{\rho}) \exp(i\beta_\lambda z)$ with

$$U_\lambda(\vec{\rho}) = \sum_{j=0}^{N-1} \alpha_j^{(\lambda)} e^{-|\vec{\rho}-\vec{\rho}_j|^2/2\rho_0^2} \tag{1}$$

describes a traveling mode of the MC fiber. The coefficients $\alpha_j^{(\lambda)}$ and the modewidth parameter ρ_0 are variational parameters that are chosen so as to maximize the largest propagation constant β_λ corresponding to the lowest-order guided mode.

This procedure, when supplemented with symmetry considerations as described fully in Ref. 3, enables us to compute the characteristics of all 7 guided modes for a symmetrical 7-core fiber, in which a core is surrounded by symmetrically located 6 other cores. For the underlying hexagonal geometry, the modes can be classifed into two sets. The first set has 5 modes, labeled $\lambda = 1, .., 5$, in each of which the central core is completely unexcited, while the surrounding outer cores are excited with amplitudes

$$\alpha_j^{(\lambda)} = e^{i2\pi\lambda j/6} \quad \text{for} \quad j = 1, .., 6, \tag{2}$$

which are the complex amplitudes at the core centers corresponding to λ complete waves around the circle containing those centers. These modes do not all have different propagation constants. In fact, a little reflection shows that the $\lambda = 1$ and $\lambda = 5$ (and similarly $\lambda = 2$ and 4) modes have amplitudes on each core that are merely complex conjugates of each other. Under complex conjugation, a forward traveling mode merely reverses direction without a change in the propagation constant, so it is evident that $\lambda = 1, 5$ and $\lambda = 2, 4$ define degenerate pairs of modes. Thus there are only five distinct propagation constants, $\beta_\pm, \beta_1, \beta_2$, and β_3. The complex-conjugation symmetry is actually equivalent to the reflection symmetry of the core geometry about any diameter passing through oppositely located cores. The second

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

set has two modes, labeled $\pm$, in each of which the outer cores are all excited with the same amplitude and phase, while the central inner core has an amplitude that is $|\alpha_\pm|$ (computed in detail in Paper I) times that of any outer core and has in-phase (for the $+$ mode) or out-of-phase (for the $-$ mode) excitation relative to any outer core's.

All of the mode propagation constants can be computed by means of the ratio

$$\beta_\lambda^2 \equiv \frac{\int d^2\vec{\rho}\, U_\lambda^*(\vec{\rho})[\nabla_T^2 + k^2 n^2(\vec{\rho})]f(\vec{\rho})}{\int d^2\vec{\rho}\, U_\lambda^*(\vec{\rho}) U_\lambda(\vec{\rho})}, \tag{3}$$

which can be computed in terms of Gaussian integrals. It is the $+$ mode that, of all 7 modes, has the softest core-to-core complex-amplitude variation in the transverse plane, and so has the largest longitudinal wavevector or propagation constant β_+. The final variational parameter ρ_0 was determined numerically; it was adjusted until the largest eigenvalue β_+ attained its maximum value.

This finishes our brief review of the variational method.

III. Coupling Atmospherically Corrupted Waves: Application to Ground-Based Interferometry

As we have remarked earlier, use of a fiber for beam coupling, transport, and recombination in an interferometer is highly desirable for many reasons, two of which are its mechanical flexibility and seeing independence of fringe visibility in interference of two SM fiber beams. The use of a MC fiber has been motivated by the need to enhance the low coupling efficiency of a standard SM fiber, without compromising the perfect spatial coherence of its guided mode. We consider here coupling wavefronts emitted by a distant point source which in arriving on the ground have propagated through a turbulent atmosphere characterized by rapid refractive-index fluctuations and so have become highly aberrated. We shall consider this problem for a wide variety of atmospheric turbulence conditions, and show that the coupling efficiency can improve many-fold with the use of a MC fiber, rather than the standard SCSM fiber, depending on the severity of turbulence.

Imagine a distant point source in the sky that emits a spherical wave of a single frequency $\omega = ck$. By the time the wave arrives at the top of the terrestrial atmosphere, it is essentially a plane wave over large transverse distances with a constant electric field E_0, say of amplitude 1. Were it not for the inhomogeneities of the atmosphere, the wave at the ground would have been just the phase factor $\exp(in_0\omega L/c)$ corresponding to an atmospheric optical depth $n_0 L$. But due to turbulence-induced refractive-index fluctuations, the wavefront can only statistically be described. In terms of the Fried parameter, r_0, which describes the extent of spatial coherence in an atmospherically corrupted wavefront, the second-order field correlation function $\Gamma_0(|\vec{r} - \vec{r}'|)$ in the wavefront can be written down in the form

$$\begin{aligned}\Gamma_0(|\vec{r} - \vec{r}'|) &\equiv \langle E_0(\vec{r}) E_0^*(\vec{r}')\rangle \\ &= \langle e^{i\phi(\vec{r})} e^{-i\phi(\vec{r}')}\rangle \\ &= e^{-3.44(|\vec{r}-\vec{r}'|/r_0)^{5/3}[1-(|\vec{r}-\vec{r}'|/D)^{1/3}]}.\end{aligned} \tag{4}$$

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

The power coupling fraction η into guided modes of the MC fiber can be expressed in terms of the overlap integrals

$$O_\lambda \equiv \int E(\vec{\rho}) U_\lambda^*(\vec{\rho}) d^2\vec{\rho} \tag{5}$$

between the guided-mode profiles $U_\lambda(\vec{\rho})$ and the focal-plane field profile $E(\vec{\rho})$ and the mode normalizations

$$N_\lambda \equiv \int |U(\vec{\rho})|^2 d^2\vec{\rho} \tag{6}$$

as

$$\eta = \sum_{\lambda=\pm,1}^{5} |O_\lambda|^2 / N_\lambda. \tag{7}$$

Since the focal-plane field is merely a Fourier transform of the pupil-plane field times the pupil function, i.e.

$$E(\vec{\rho}) = \frac{k}{2\pi i f} \int P(\rho'; D/2) E_0(\vec{\rho}') e^{-i(k/f)\vec{\rho}\cdot\vec{\rho}'} d^2\vec{\rho}', \tag{8}$$

(f: focal length and P is 1 for $|\vec{\rho}'| < D/2$ and 0 elsewhere), one can express the overlap integral, written in (5) in the focal plane, in the pupil plane according to Parseval's theorem. That latter form involves the Fourier transform of the fiber mode U_λ, obtained by propagating it backwards to the pupil plane, which can easily be done for a superposition of identical but displaced Gaussian forms. When this is substituted into (7), the coupling efficiency η is obtained for a given realization of the pupil wavefront. Of course, since the pupil wavefront can only statistically be specified, the coupling efficiency η too can only statistically determined. By taking the average of the resulting expresion over atmospheric fluctutations, the following average coupling efficiency is obtained:

$$\langle\eta\rangle = 16 \left(\frac{\rho_0}{a}\right)^2 \tilde{a}^2 \int_{\rho'<1/2} d^2\vec{\rho}' \int_{\rho''<1/2} d^2\vec{\rho}'' \Gamma_0(D|\vec{\rho}' - \vec{\rho}''|) e^{-2\pi^2\left(\frac{\rho_0}{a}\right)^2 \tilde{a}^2(\rho'^2 + \rho''^2)}$$

$$\times \sum_\lambda \frac{1}{(N_\lambda/\pi\rho_0^2)} \sum_{j=0}^{6} \sum_{\ell=0}^{6} \alpha_j^{(\lambda)} \alpha_\ell^{*(\lambda)} e^{i2\pi\left(\frac{R}{a}\right)\tilde{a}(\vec{\rho}_j\cdot\vec{\rho}' - \vec{\rho}_\ell\cdot\vec{\rho}'')}, \tag{9}$$

in which $\tilde{a}$ is the ratio of core radius a and the characteristic size of the diffraction-limited Airy disk, $R_D \equiv 2\pi f/(kD)$, i.e.,

$$\tilde{a} = a/R_D = akD/(2\pi f). \tag{10}$$

It is form (9) that we numerically computed for different $R/a, V, a/R_D$ values under varying conditions of atmospheric turbulence characterized by different D/r_0 values.

In Figs. 2 and 3, we have plotted the average power coupling efficiency $\langle\eta\rangle$ as a function of the core-radius-to-Airy-disk-radius ratio for $D/r_0 = 5$ and 10, respectively. In each figure, two V values, 2 and 2.4, were used and the core-core separation was maintained at 3 times each core radius. For comparison, the standard SCSM fiber result has also been plotted for each V value; this was obtained directly by limiting

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

the λ as well as the j and ℓ sums in (10) to one term and using the relation $\rho_0/a = 1/\sqrt{2\ln V}$ valid for the standard SCSM fiber.[7] The improvement is already dramatic for $D/r_0 = 5$, for which the MC fiber numbers are 33.6% and 29% for $V = 2$ and 4, respectively, as opposed to the SCSM fiber numbers of about 11%. The factor 3 improvement with $D/r_0 = 5$ is exceeded by the factor 5.5 improvement when D/r_0 is 10; the numbers for the MC fiber are 11% and 10%, while the SCSM fiber comes in at about 2%. There is a clear trend: the worse the distortions of the wavefront, the better the relative coupling efficiency of the MC fiber. This is because each individual core of a MC fiber is exposed to a rather small portion of the wavefront with correspondingly high coherence relative to the single core of the standard SCSM fiber which sees almost the whole wavefront with much lower coherence. Since the coupling to the cores of a fiber is a coherent overlap process, the MC fiber is guaranteed to couple more strongly. Note the implications of this result for ground-based interferometry. Without the use of adaptive optics, with 1-meter-aperture telescopes operating under a typical r_0 of about 10 *cms* at optical wavelengths, substituting MC fibers for SCSM fibers would enable the interferometer to image sources that are about 5 times fainter. This would represent a gain in limiting visual magnitude of about 2 over that achievable with conventional fibers.

There is another extremely important observation to make here. Note that for optimal coupling, the ratio of the core radius to the radius of the diffraction-limited Airy pattern is roughly between 0.5 and 0.6, *regardless* of the strength of turbulence as represented by the D/r_0 ratio. For a fixed pupil diameter D, this implies independence of the optimal core radius from the conditions of turbulence, so that the same MC fiber will optimally couple for varying r_0 values. Such robustness of the MC-fiber design arises fundamentally from the fact that in the image plane, the structure (although not the magnitude) of field correlations, or in other words spatial coherence, is independent of any variations of r_0, depending only on the pupil diameter.[8] Fiber coupling, being a coherent overlap process as noted above, should therefore remain optimized for a given aperture diameter, independent of r_0 or the degree of any adaptive corrections or any other source of pupil-plane aberrations. It is worth noting that this result is valid for any coherent coupling process, including coupling into a standard SCSM fiber.

IV. Propagation of Guided Field and Evolution of Core-Core Coherence

Having addressed the coupling process, we now turn to the problem of propagation of the guided field inside a MC fiber. This is an important consideration, since at some point down the fiber length, the cores have to be physically separated to form 7 individual SCSM fibers carrying 7 sub-beams. For ease of calibration of the fringe visibility, it is useful to ensure that all sub-beams have roughly the same strength, even though at the initial coupling in the focal plane, the central core would clearly see a larger field and so carry a stronger sub-beam than the outer ones. But as the field traverses the fiber, evanescent fields cause the overall field configuration to evolve. An example of this evolution in which a plane wave was focused and coupled into the fiber was covered in Ref. 3.

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

Here, we discuss the scenario in which atmospherically aberrated wavefronts are focused into a 7-core fiber. When aberrated waves are involved, a lack of azimuthal symmetry of the field being coupled in means that all 7 guided modes of the MC fiber will in general get excited, not just the $\pm$ modes. However, when a wavefront is statistically described with a point-to-point structure function that depends only upon the distance between the points, then the azimuthal symmetry of the pupil again implies certain decorrelations in the excited modes of the MC fiber. It is this sort of aberrated wavefront that is of concern to us here.

For an amplitude g_λ of excitation of mode λ, the jth core excitation amplitude $a_j(z)$ after distance z of propagation is given by the relation

$$a_j(z) = \sum_\lambda g_\lambda a_j^{(\lambda)} e^{i\beta_\lambda z}, \tag{11}$$

where $g_\lambda = O_\lambda/N_\lambda$. Since g_λ is proportional to the overlap integral O_λ between the pupil-plane field and the mode field distribution propagated backwards to the pupil, we can express the core-core correlation matrix elements

$$c_{j\ell}(z) \equiv \langle a_j(z) a_\ell^*(z) \rangle \tag{12}$$

in terms of the pupil-plane field correlations. The final result of this analysis is

$$c_{j\ell}(z) = \sum_\lambda \sum_{\lambda'} a_j^{(\lambda)} a_\ell^{(\lambda')*} e^{i(\beta_\lambda - \beta_{\lambda'})z} \langle g_\lambda g_{\lambda'}^* \rangle, \tag{13}$$

where the mode-mode correlations are given by the expression

$$\langle g_\lambda g_{\lambda'}^* \rangle = \frac{4}{(N_\lambda/\pi\rho_0^2)(N_{\lambda'}/\pi\rho_0^2)} \frac{(D/r_0)^4}{(2\pi f/kr_0^2)^2} \sum_{n=0}^{6} \sum_{m=0}^{6} \alpha_n^{(\lambda)*} \alpha_m^{(\lambda')} \int_{\rho'<1/2} d^2\vec{\rho}\,' \int_{\rho''<1/2} d^2\vec{\rho}\,''$$

$$\times \tilde{\Gamma}_0(|\vec{\rho}\,' - \vec{\rho}\,''|) e^{-2\pi^2 \left(\frac{\rho_0}{a}\right)^2 \tilde{a}^2 (\rho'^2 + \rho''^2)} e^{i2\pi\left(\frac{R}{a}\right)\tilde{a}(\vec{\rho}_n \cdot \vec{\rho}\,' - \vec{\rho}_m \cdot \vec{\rho}\,'')}, \tag{14}$$

Given the form (14) of mode-mode correlations, it can be shown that for $\lambda, \lambda' = 1, .., 5$, the modes are uncorrelated, i.e.,

$$\langle g_\lambda g_{\lambda'}^* \rangle = 0 \quad \text{for} \quad \lambda \neq \lambda' = 1, .., 5. \tag{15}$$

It can be further shown that the correlations between the $\lambda = 1, .., 5$ modes and the $\pm$ modes also vanish. In fact, the only correlation that survives is that between the $\pm$ modes. In view of such decorrelations, Eq. (13) shows clearly that any dependence of the core-core correlations on the propagation distance z must come only from the $\pm$-mode correlations. The core-core correlations are thus characterized by a single sinusoidal function and are therefore perfectly periodic in z, with a period equal to $2\pi/|\beta_+ - \beta_-|$. Thus, when the aberrated wave has an isotropic and homogeneous statistical form, the variations of core-core correlations with propagation distance are characterized by one and the same period. Note that

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

by setting $j = \ell$ in (13), we can obtain the rms excitation of a particular core j, while the normalized correlation of fields in the inner and an outer core is obtained by taking the ratio

$$\gamma_{01} = \frac{c_{01}}{\sqrt{c_{00}c_{11}}}. \tag{16}$$

In Fig. 4, we have displayed the evolution of the rms excitation of any outer core and of the inner core for $D/r_0 = 10$. The propagation distance in these figures is in units of $2n_{cl}/(k\Delta n^2)$, which is equal to the vacuum wavelength times $n_{cl}/(\pi\Delta n^2)$. Two values of R/a, namely 2.5 and 3.5, were used while the core radius was fixed at $0.6R_D$ and V at 2. For the case of closer cores, clearly the cross-talk between the cores is higher and the exchange of energy between the inner and outer cores happens over a shorter distance. Note also that for the farther-core case although the initial excitations of the inner and outer cores are more disparate, with propagation they approach each other more closely. In fact, they become approximately the same after a half integral number of periods of propagation. In order to design a MC fiber with the capability of producing equal-strength – or nearly so – sub-beams at its output, we need to separate the cores of the fiber at about half the period (at which the first closest approach of inner-outer core excitations occurs) into individual SCSM fibers, each carrying a sub-beam. The separation process must take place adiabatically along the propagation direction, so as to prevent leakage of energy into radiation modes of the intermediate core geometries through which the MC-fiber core geometry must evolve.

In Fig. 5, we have plotted the normalized correlation γ_{01} between the fields carried by the inner core and an outer core of the MC fiber, as a function of the propagation distance in the same units as in the earlier figures. Since for $D/r_0 = 5$ the image-plane correlations are stronger[8] than for $D/r_0 = 10$, the inner-outer core fields are expected to be initially more strongly correlated in the former case. What is remarkable, however, is that as the core excitations approach each other (see Fig. 4) with propagation, that initial correlation improves significantly. The improvement is by a factor of 2 for $R/a = 2.5$ and about 7 for $R/a = 3.5$, both at $D/r_0 = 5$. For $D/r_0 = 10$, these improvement factors are understandably lower, by a factor 2 or so at best. Since the separation of the MC fiber into 7 individual SCSM fibers is designed to occur when the core excitations are the closest, the sub-beams generated in this manner will also have the highest correlations possible. Enhanced correlations are of great value for interferometry, since they represent reduced mutual phase and amplitude fluctuations of the sub-beams. As an example, consider the normalized correlation of about 0.45 possible with $D/r_0 = 5$ and $R/a = 3.5$. If all of the decorrelation implied by this value were to be due to phase fluctuations, then this would mean a root-mean-square (rms) mutual phase fluctuation $\Delta\phi_{rms}$ of the core fields given by the following expression valid for Gaussian statistics:

$$\gamma_{01} = 0.45 = e^{-(1/2)\Delta\phi_{rms}^2},$$

which amounts to $\Delta\phi_{rms}$ about 1.26 rad, or about a fifth of a wave. These considerations suggest very minimal requirements of phase and amplitude compensations to be implemented in each arm of the interferometer using MC fibers, even under conditions of high turbulence.

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

As we have noted before, the fundamental period of evolution of the core-core excitations and correlations is fixed by the the MC fiber, being equal to $2\pi/|\beta_+ - \beta_-|$, and not by conditions of turbulence, i.e. not by the D/r_0 ratio. This is clearly exhibited in Figs. 4 and 5. A similar kind of independence from atmospheric conditions, although arising for entirely different reasons, was noted in Sec. III for the ratio a/R_D needed for optimal coupling. The insensitivity of a/R_D that optimizes the initial coupling fraction and of the period of the evolution of the core excitations and correlations with respect to atmospheric and other aberration conditions is clearly of utmost importance to the design of a MC-fiber based interferometer.

V. Conclusions

In this paper, we have reporte on the development of a variational scalar-field approach to the computation of the guided field modes of a multiple-core fiber and the use of the resulting mode characteristics to demonstrate large theoretical enhancements in coupling efficiencies when MC fibers rather than standard SCSM fibers are used. The specific MC fiber we have considered in detail is one with 7 cores. The enhancement can be many fold; for $D = 5r_0$-sized apertures, the improvement factor is about 3 while for $D = 10r_0$, that factor rises to 5.5. Such efficiency gains represent impressive limiting visual magnitudes gains, upto about 2. Furthermore, the core radius at which maximum coupling takes place is independent of the conditions of turbulence, provided the pupil diameter is kept fixed.

We also computed the evolution of the core excitations and coherences as a function of the propagation distance in the MC fiber. Although initially energy is concentrated in the inner core, with propagation evanescent fields cause a periodic transfer of average energy between the inner and outer cores, even when the input wavefront is corrupted by atmospheric turbulence. This periodicity as well as the value of the period is set by the design of the MC fiber, and is independent of the atmospheric or other aberration conditions. The inner-outer-core mutual coherence has also the same periodic evolution, and in fact when the average core excitations have the smallest difference, at half-integral periods along z, the mutual coherences are also at their peak. Thus, the optimal value of the distance at which the MC fiber is to transformed into single-core fiber outputs is the same from the viewpoints of both the relative core-core excitation and the mutual core-core coherences. Furthermore, this optimal criterion, like the value of the core radius for maximum coupling, is entirely independent of the strength of turbulence, which is a highly desirable feature of our proposed design.

Acknowledgments

FOCUSING LIGHT INTO A MULTIPLE-CORE FIBER

Sudhakar Prasad

References

1. S. Shaklan and F. Roddier, "Single-mode fiber optics in a long-baseline interferometer," Appl. Opt. **26**, 2159-2163 (1987).

2. P. Connes, S. Shaklan, and F. Roddier, "A fiber-linked ground-based array," in *Proc. Joint NOAO-ESO Workshop on High-Resolution Imaging from the Ground Using Interferometric Techniques, Oracle, Arizona*, ed. J. Goad (NOAO, Tucson, Arizona, 1987), 165-168.

3. S. Prasad, "Focusing light into a multiple-core fiber: I. Theory," to be published.

4. S. Prasad, "Focusing light into a multiple-core fiber: II. Application to ground-based interferometry," to be published.

5. H. McAlister, W. Bagnulo, T. ten Brummelaar, W. Hartkopf, N. Turner, A. Garrison, W. Robinson, and S. Ridgeway, "The CHARA Array," in *Amplitude and Intensity Spatial Interferometry II*, J. Breckenridge, ed., Proc. SPIE 2200, 129-139 (1994).

6. D. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures," J. Opt. Soc. Am. **56**, 1372-1380 (1966).

7. A. Snyder and J. Love, *Optical Waveguide Theory*, (Chapman and Hall, London, 1983), Chap. 15.

8. S. Prasad, "Statistical field correlations in the image plane of a ground-based telescope," Opt. Commun. **104**, 251-258 (1994).

Figure Captions

Fig. 1. The hexagonal geometry of a symmetrical 7−core fiber.

Fig. 2. A plot of the average power coupling efficiency, defined as the ratio of average power coupled into the guided modes of the MC fiber and that incident in the pupil, as a function of the core radius. The incident wavefront was tilt-corrected over an aperture of size $D = 5r_0$. The upper pair of curves (full lines) represents the MC-fiber coupling efficiency while the lower pair of curves (dashed lines) represents the SCSM fiber coupling. In each pair, the upper curve corresponds to $V = 2$ and the lower to $V = 2.4$.

Fig. 3. Same as Fig. 2, except that $D = 10r_0$.

Fig. 4. Periodic exchange of excitation between the inner and outer cores of the MC fiber, for $V = 2.0$, $a/R_d = 0.6$, and $D = 10r_0$. The pairs of curves shown full and dotted each correspond to two different core-core separations, namely, $R/a = 2.5, 3.5$, respectively. The inner-core excitation is always higher initially.

Fig. 5. Periodic evolution of the normalized correlation between the inner and an outer core of the MC fiber, for $V = 2.0$ and $a/R_d = 0.6$. Two different D/r_0 ratios and R/a are represented; $D = 5r_0$ for the full ($R/a = 2.5$) and the dotted ($R/a = 2.5$) curves, while $D = 10r_0$ for the dashed ($R/a = 2.5$) and the dot-dashed ($R/a = 2.5$) curves.

Fig. 1

COUPLING FOR 7-CORE VS. 1-CORE FIBER FOR $D/r_0 = 5$

Fig. 2

COUPLING FOR 7-CORE VS. 1-CORE FIBER FOR $D/r_0 = 10$

Fig. 3

INNER VS. OUTER CORE EXCITATION: V=2, $a/R_D=0.6$, $D/r_0=10$

R/a = 2.5

R/a = 3.5

RMS EXCITATION

PROP. DISTANCE z

Fig. 4

INNER-OUTER CORE CORRELATION: V=2, $a/R_D = 0.6$
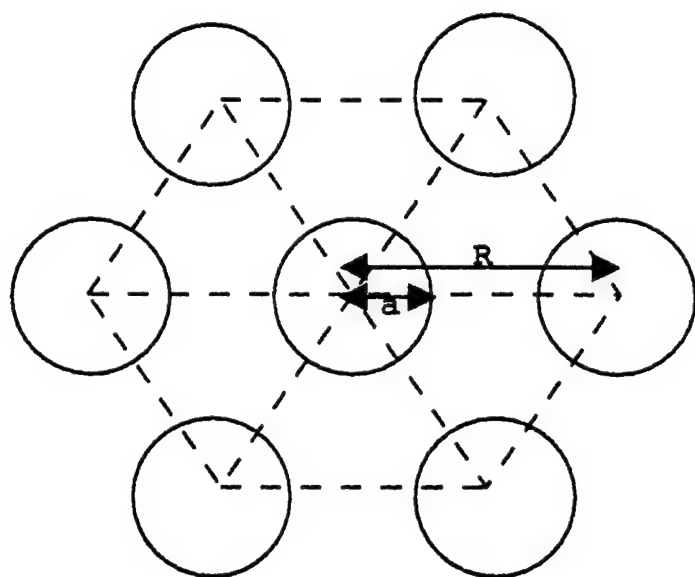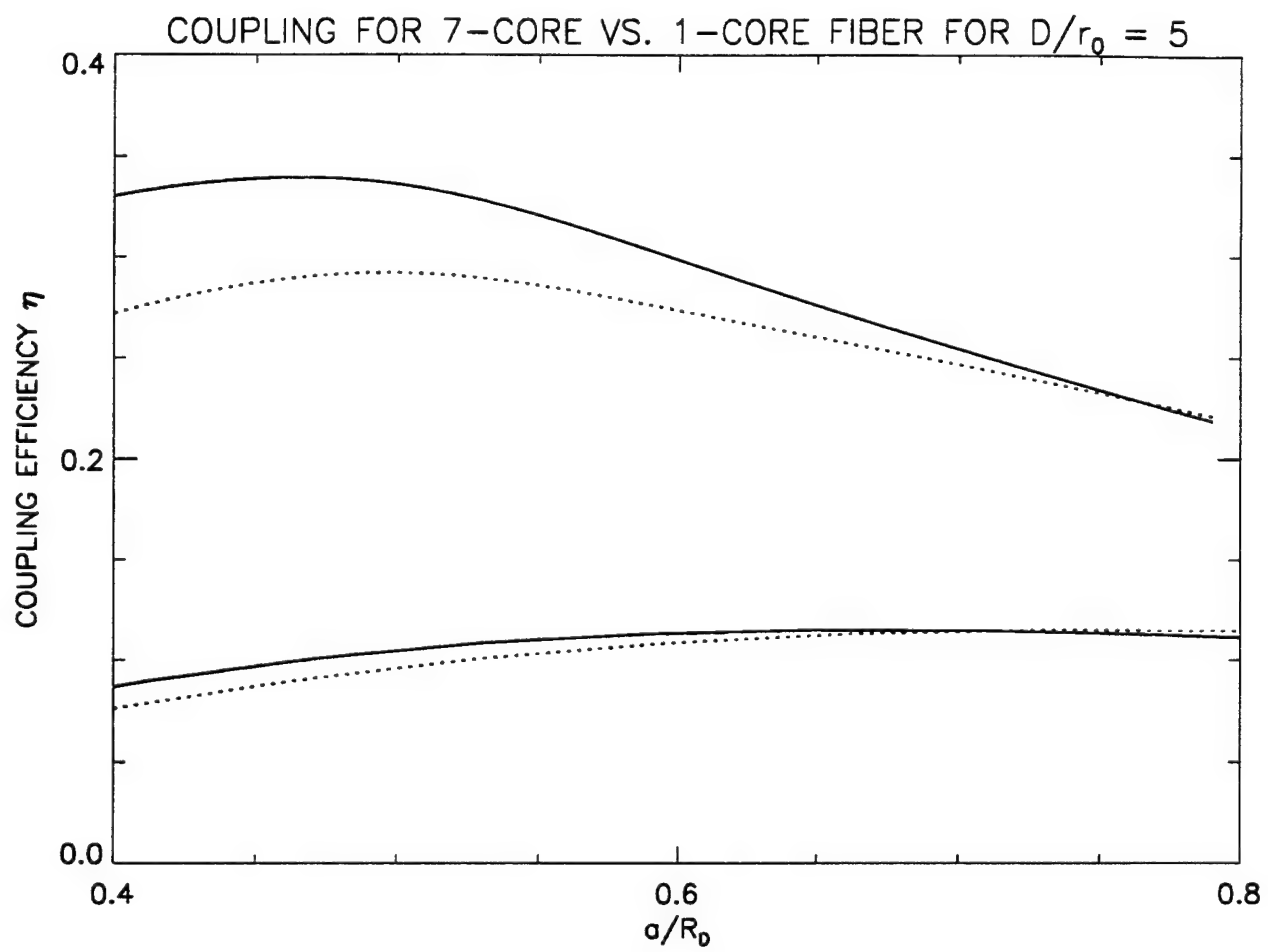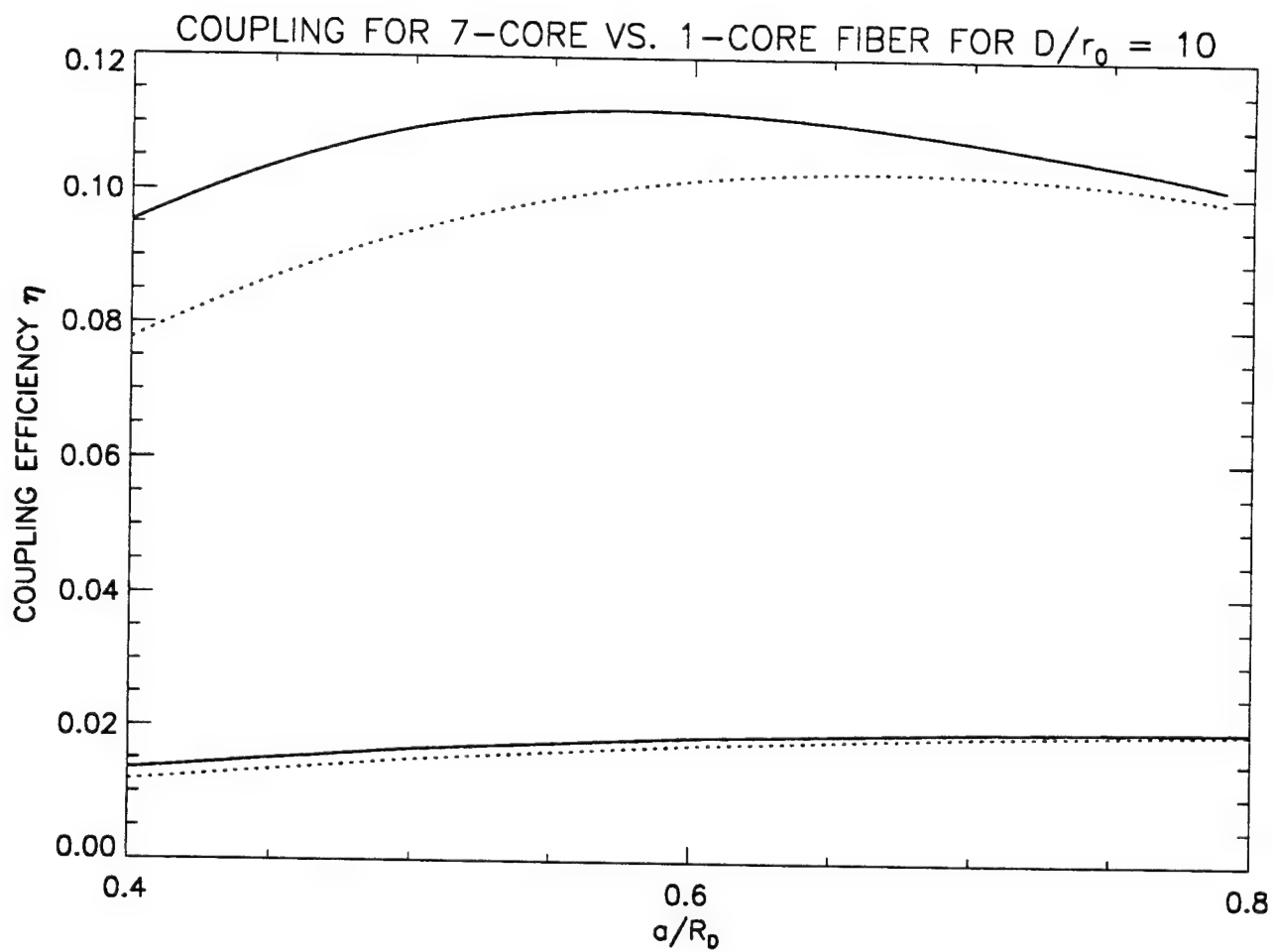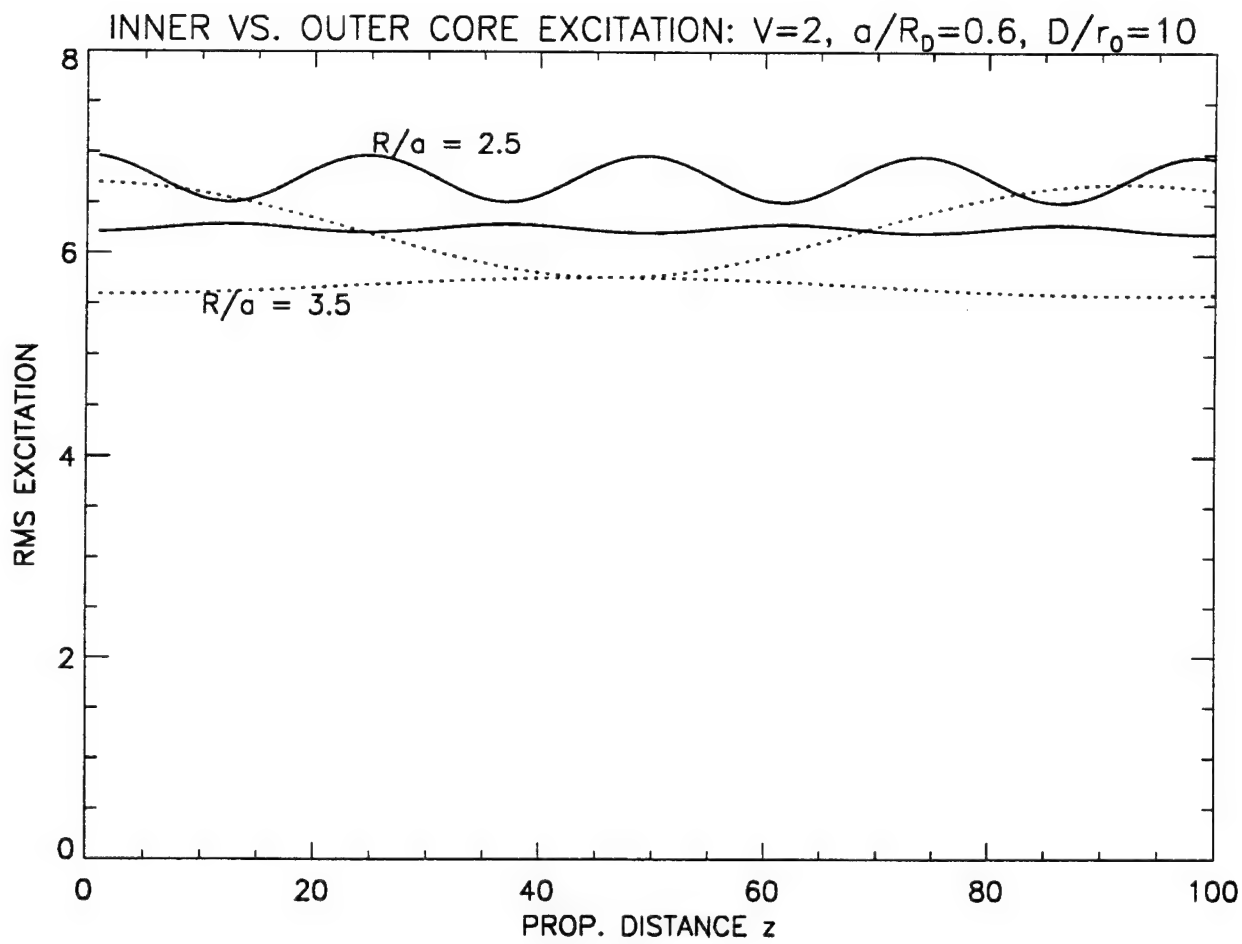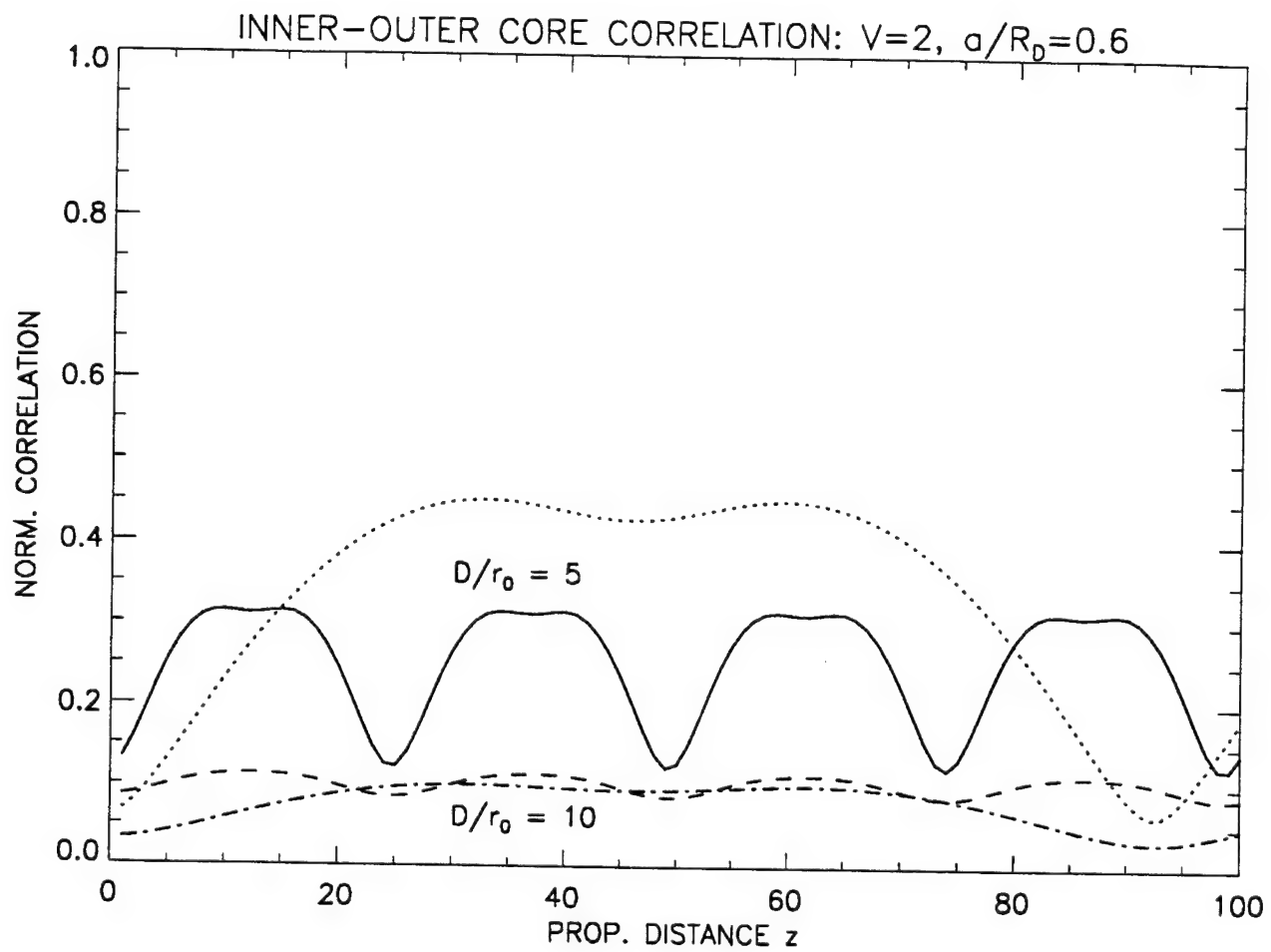
Fig. 5

STATIC AND DYNAMIC GRAPH EMBEDDING FOR PARALLEL PROGRAMMING

Dr. Mark Purtill
Assistant Professor
Department of Mathematics

Texas A&M University-Kingsville
Campus Box 172
Kingsville TX 78363

STATIC AND DYNAMIC GRAPH EMBEDDING FOR PARALLEL PROGRAMMING

Dr. Mark Purtill
Assistant Professor
Department of Mathematics
Texas A&M University-Kingsville

Abstract

Many problems in computational physics can be modeled as a computation on a graph or network, for example, particle-in-cell plasma simulations. Solving these problems on a massively parallel processor is becoming increasingly important. As many parallel processors can also be modeled as graphs, this gives rise to the problem of efficiently embedding the "computation" graph into the "processor" graph. We survey some of the work done in this area.

For best performance, this "static" embedding must be supplemented by a dynamic re-embedding (or load balancing) as load in a time shared parallel processor varies, or the amount of computation to be done at each vertex of the computation graph changes. (The latter is a particular concern in PIC plasma simulations.) We give a new algorithm for this re-embedding based on network flows, and discuss an implementation of this algorithm and preliminary results from a simulated problem on real parallel hardware.

STATIC AND DYNAMIC GRAPH EMBEDDING FOR PARALLEL PROGRAMMING

Dr. Mark Purtill

1   Introduction

Many problems in computational physics can be modeled as a computation on a graph or network, for example, particle-in-cell plasma simulations. Solving these problems on a massively parallel processor is becoming increasingly important. As many parallel processors can also be modeled as graphs, this gives rise to the problem of efficiently embedding the "computation" graph into the "processor" graph.

For best performance, this "static" embedding must be supplemented by a dynamic re-embedding as load in a time shared parallel processor varies, or the amount of computation to be done at each vertex of the computation graph changes. (The latter is a particular concern in PIC plasma simulations.)

The paper is organized as follows. In the next section, we define the relevant graph theory terminology and formulate the problem. This section concludes with an example from computational physics. Section 3 discusses various approaches to the "static" embedding problem, while section 4 discusses our algorithm for the re-embedding problem. In section 5 we discuss the implementation, followed by conclusions and suggestions for future work.

2   Graph Computation

2.1   Graph Terminology

A *graph* is an abstract combinatorial structure. A graph consists of a set of *vertices* (singular vertex), some pairs of which are connected by *edges*.[1] Thus, a graph G can be represented as a pair $(V(G), E(G)) = (V, E)$, where V is the set of vertices, and E is the set of edges. The set of edges can be thought of as a subset of the set of all pairs of distinct vertices $\binom{V}{2}$.[2]

We represent graphs by drawings where the vertices are points on the plane, and the edges lines or curves connecting the two vertices, but the geometric information in the drawing is *not* part of the graph. For instance, the two diagrams in figure 1 represent the same graph. If two vertices u and v are connected by an edge, they are said to be *adjacent*, and we write $u \sim v$. The collection of vertices adjacent to a given

---

[1] Graph theory terminology varies somewhat. Vertices are also called *nodes* or *points*; edges are also called *arcs*.

[2] Here we do not allow multiple edges between the same pair of vertices, nor do we allow an edge connecting an edge to itself.
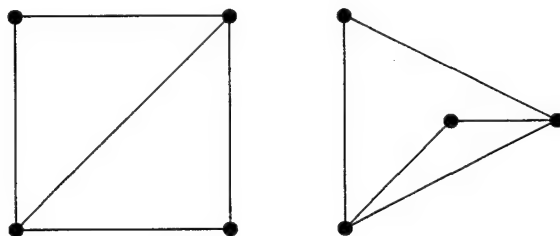
Figure 1: Two diagrams representing the same graph

vertex v is called the *neighborhood* of the vertex v; the number of vertices adjacent to v is called the *degree* of v.

A *path* is a sequence of vertices $v_0, v_1, \ldots, v_k$ such that each pair $\{v_i, v_{i+1}\}$ is an edge. We also sometimes consider the path to be the sequence of edges. If $k = 0$, so there are no edges, the path is said to be *trivial*. The *length* of a path is the number of *edges* in it (so a trivial path has length 0).

2.2 Graph Computation

In this paper, we will be considering computations on a graph. We assume that some information is stored at each vertex; the computation proceeds as follows:

1. at each vertex v, information from v is sent to the vertices adjacent to v; then

2. at each vertex v, computation is done at v using the information at v along with information sent to v from its neighbors in the first step.

We are most interested in doing this algorithm in parallel: each vertex can do its computation (the second step) in parallel, and, depending on the parallel model, possibly some of the communications in parallel as well.

2.3 Parallel Machines as Graphs

Many distributed-memory parallel processors can be considered as graphs. Each vertex represents one processor, and each edge would represent a dedicated communications channel between the two processors.[3] To avoid confusion, we will refer to the vertices of a parallel processor graph as *processors* and to the edges as *communications links* or just links.

For instance, a graph of an eight dimensional hypercube is seen in figure 2. To run a computation to be done on a graph C on a parallel processor graph P, we must decide on an *embedding* of the graph C into the graph P, that is:

- for each vertex $v \in V(G)$, decide which processor $\mathbf{p}(v) \in V(P)$ will do the computation for v; and

- for each edge $e = \{u, v\} \in V(G)$, find a path (sequence of links) $\mathbf{p}(e)$ in P from $\mathbf{p}(u)$ to $\mathbf{p}(v)$. (If $\mathbf{p}(u) = \mathbf{p}(v)$, we can choose $\mathbf{p}(e)$ to be trivial.)

There are a variety of ways we can measure the goodness of such an embedding, such as:

dilation the length of the longest path $\mathbf{p}(e)$;

---

[3] Perhaps the theory of *hypergraphs* could be used to handle communications channels such as buses. In a hypergraph, edges may have more than two vertices, but the theory is less well developed.
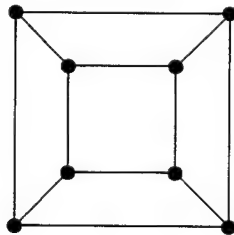
Figure 2: An eight-node hypercube-architecture parallel processor

balance the ratio of the number of [total weight of] the vertices on the processor with the most vertices [most total weight of vertices] to that with the least vertices [least total weight].

For our purposes, the embedding which produces the shortest run time for a step of computation would be best. The exact mathematical description of this would depend on the details of the parallel machine. (For instance, whether communications can be done simultaneously by a given processor; whether sends and receives are blocking or nonblocking; whether multiple sends, multiple receives, or both can be done in parallel, and so on.)

Suppose our parallel machine is such that messages can be received asynchronously (that is, without using the main processor), and the problem is such that we will not swamp any processor with incoming messages. Then the total time for a given processor p to complete a step will be approximately:[4]

$$\sum_v [\![\mathbf{p}(v) = p]\!]w(v) + \max\Big(\sum_{u \sim v}[\![\mathbf{p}(u) = p]\!][\![\mathbf{p}(v) \neq p]\!]w(u,v),$$

$$\max_{q \sim p}\sum_{u \sim v}[\![\mathbf{p}(u) = q]\!][\![\mathbf{p}(v) \neq q]\!]w(u,v)\Big)$$

where $w(v)$ is the amount of computation that must be done at vertex v, and $w(u,v)$ is the amount of communications from u to v (and is assumed to be 0 if there is no edge $\{u,v\}$ in G). The first term represents the time necessary to do the computation, and the second that for communications, which is the maximum of (the time necessary to send communications from processor p) and (the time to receive communications from p's neighbors). (The last expression is the total communications time at each neighboring vertex q; clever arrangement of the communications could improve this, depending on the exact details of the computation.)

This expression assumes all processors and all communications channels have equal speed; this is not usually a good assumption. It is obviously a bad assumption in the case of a heterogeneous machine, but even with a homogeneous machine which is time-shared, different jobs running on different processors could result in inhomogeneous effective speeds for the different processors.

The total time to complete a step would be the maximum value of this expression over all processors p. For different models of parallel communications, and depending on the details of the problem, various of the maxs could become $\sum$s and vice versa.

2.4 The Minimum Communications Graph Embedding Problem

Because of the differences between different parallel machines, as well as the complicated nature of the "correct" expression to be minimized, we will consider on the following simplified problem (which should still produce reasonably fast running times).

> Given a computation graph C and a processor graph P, both possibly weighted on edges or nodes, find an embedding of C into P so that an equal weight of nodes of C are mapped to each node of P, and communications is minimized (that is, with a minimal weight of edges of C mapped to edges of P).

This problem is NP-complete. It is obviously in NP (as a decision problem) since the cost of a partition can be computed in polynomial time.

To see the problem is NP-hard, let C be any graph and P be a graph with two vertices and one edge connecting them. Setting all the weights to 1, our problem reduces to partitioning a graph into two equal parts with a minimal number of edges crossing, which is well known to be NP-complete. (See [6].)

2.5 Examples from Computational Physics

Many problems involving numeric solving of partial differential equations can be cast as graph computation. However, the motivating example for this work is a particle-in-cell method for numerical computation of medium density plasma.

---

[4] Here we are using the notation that for any statement χ, the expression $[\![\chi]\!]$ is 1 if χ is true and 0 if χ is false.

In a plasma, there are magnetic and electrical fields interacting with charged particles moving in space; the interaction is governed by Maxwell's equations (for the fields) and the Lorentz equations (for the motions of the particles). In the particle-in-cell model, space is divided into cubes (cells) by a grid. The fields are kept track of on the grid points (or, in more sophisticated codes, on faces or edges of the cubes, with different field components having different locations). The particles, on the other hand, are allowed to move freely; they are kept track of by which cell they are in, then the coordinates of the particle within the cell. For more details, see for instance Birdsall and Langdon [3]. Minor [17] gives details of a more sophisticated PIC code.

The computation graphs we get from PIC codes will have a nice structure. Consider the case where all field components are kept track of at the grid points; then the graph will look like the gird, with the addition of a point at the center of each cell connected to the vertices of that cell as well as the center point of adjacent cells. Thus, there would be two types of vertices: one representing the grid points, and another representing the cells. In each case, those vertices not on or near the boundary would have degree 14. However, this does not limit us to "boxes"; see, for example, figure 3. (Here we only show one grid for clarity). This is a small example of a cavity which is of interest to plamsa physicists.

Because of the movement of particles, PIC simulations have an additional feature: the amount of work to be done at the vertices representing the cells will change over time as particles move in the system. The amount of computation needed will be (in general) will be unpredicatable, so a dynamic load balancing scheme will be required for good performance.

3 Static Embedding

3.1 Heuristic Approaches

As the Minimum Communications Graph Embedding problem is NP-hard, we do not expect to find an exact algorithm to solve the problem that will run in subexponential time. Therefore, most work has concentrated on heuristic approaches. In this section, we will survey some of the approaches suggested. Other surveys can be found in William [25] and Hendrickson and Leland [9].

Most heuristics concentrate on partitioning the graph into two equal parts with a minimal number of edges crossing the partition. In the common case of embedding into a hypercube, repeated partitioning can be used to get an embedding.

3.1.1 The Inertial Method

In the inertial method, we assume that the points of our graph represent points in space. The inertial method finds the longest axis of the points, and partitions along that.

A variant of the inertial method specifically for PIC codes, where space is divided up (particle motion being ignored or treated as a weight for different regions of space), has been investigated by Walker [24] and by Ferraro, Liewer and Decyk [4]; the latter investigated the dynamic problem as well.

3.1.2 The Spectral Method

The problem of finding a minimum cost partition can be expressed as an integer quadratic program. Let us assume for simplicity that the number of vertices of G is even and all node weights are 1, so $\#V_1 = \#V_2$. We also assume that all the edge weights are 1.

The condition that there be an equal number of vertices in each of the two parts of the partition is expressed by the linear equation $\sum_i x_i = 0$. The number of edges crossing the partition is given by $\frac{1}{8}\sum [i \sim j](x_i - x_j)^2$, since if $x_i = x_j$, $(x_i - x_j)^2 = 0$, while if $x_i \neq x_j$, $x_i - x_j = \pm 2$, so $(x_i - x_j) = 4$ and each edge

is counted twice (once each way). Then we have

$$\sum_{i \sim j}(x_i - x_j)^2 = \sum_{i \sim j}(x_i + x_j)^2 - 2\sum_{i \sim j}x_i x_j$$
$$= \mathbf{x}^T L \mathbf{x}$$

where L is the Laplacian Matrix of the graph, $L = [\, l_{ij} \mid i, j \in V \,]$, is defined by $l_{ii} = degree(i)$, and for $i \neq j$, $l_{ij} = 0$ if no edge connects i and j, and $l_{ij} = -w$ if an edge of weight w connects i and j.[5]

Thus, we have the integer quadratic program

(1)
$$\text{Minimize } \mathbf{x}^T L \mathbf{x}$$
$$\text{Subject to } \mathbf{1}^T \mathbf{x} = 0$$
$$x_i = \pm 1$$

where $\mathbf{1}$ is the vector of all 1s.

Of course, this is still NP-hard, but it is suggestive. Suppose we now replace the constraint the $\mathbf{x}$ be a ± 1-vector with the weaker linear constraint $\mathbf{x}^T \mathbf{x} = n$. We are left with a quadratic program, which is easy to solve:

(2)
$$\text{Minimize } \mathbf{x}^T L \mathbf{x}$$
$$\text{Subject to } \mathbf{1}^T \mathbf{x} = 0$$
$$\mathbf{x}^T \mathbf{x} = n$$

This program can be solved particularly easily by finding the second smallest eigenvalue of L, as follows: Let the eigenvalues of L be $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Then in can be shown that $\lambda_1 = 0$, and if the graph is connected, then $\lambda_1 < \lambda_2$. Furthermore, it is easy to see that if $\mathbf{x}$ is an eigenvector with eigenvalue λ_2 (weighted such that $\mathbf{x}^T \mathbf{x} = 1$), then it solves the quadratic program (2). (See, for instance, Hendrickson and Leland [10] for details of all of this.)

The spectral method finds this second largest eigenvalue and a corresponding eigenvector, and then rounds to find an approximate solution to the integer program (1). For instance, it might take the vertices corresponding to the $n/2$ smallest components of the eigenvector and assign them -1 (that is, put them in V_2), while putting the remaining (largest) $n/2$ vertices in V_1. This method was introduced by Pothen, Simon, and Liou [19]; Hendrickson and Leland extended it to quadrisection and octasection [8].

3.1.3 The Method of Kernigan and Lin

Kernigan and Lin's method [13] is really a method for locally improving a partition, although it can be made into a partition finding method by starting with a random partition (or one made with one of the other methods). First, the cost or savings resulting from moving one vertex from its current part to the other part is computed, and the current partition is saved as "best so far". Then, loop we loop over the following until (say) half of the vertices have been moved:

1. figure out the best pair of vertices to swap;

2. mark these as moved;

3. update movement costs;

4. if the new partition is cheaper than the current best so far, save the new partition as the best so far.

At the end of the loop, we replace the current partition with the one saved as best so far. This process can be repeated as long as better partitions are found, or for a fixed number of iterations. Fiduccia and Mattheyses [5] have devised a data structure that allows each loop to be carried out in linear time; Suaris and Kedem [23] extended this to quadrisection, and Hendrickson and Leland [11] to the general case (of partitioning into n parts).

---

[5] If A is the ordinary adjacency matrix of the graph, and D the diagonal matrix of degrees, then $L = D - A$.

3.2 Other Methods

Among other methods which have been suggested (mostly for the bijection problem) are: simulated annealing [14]; a more easily parallelized algorithm similar to Kernigan-Lin called Mob [20]; algorithms based on node-separators [16]; among others.

3.2.1 Multilevel Methods and the Chaco Package from Sandia

Multilevel methods have been suggested by Barnard and Simon [1] and Hendrickson and Leland [9, 11, 10, 8]. The basic idea is to coarsen the graph by contracting edges. This is done in several steps until the graph is fairly small (say, around 100 vertices). This graph is now partitioned (or embedded) as required (the method appears not to matter very much) and the graph uncoarsened. After each step of uncoarsening, a variant of the Kernigan-Lin algorithm is applied (for local improvement). This work also specifically deals with embedding (into a hypercube) the graph, rather than simply partitioning it.

Hendrickson and Leland have written a computer program, called **Chaco**, which implements the spectral, inertial, and multilevel methods. We used **Chaco** to generate starting partitions for our dynamic embedding runs.

4 Dynamic Embedding

In a dynamic embedding (often called *load balancing*) procedure, we assume that either the work to be done at each vertex of the computation graph varies, or that the speed of the processors vary (or both). This means in order to avoid imbalance, we have to move vertices of the computation graph between processors. This is harder than the static problem, since we have an additional condition: we would like to keep the communication necessary to do the movement to a minimum as much as possible. Work by Smith and Baker [21] and others suggests that dynamic load balancing is vital to getting good performance from massively parallel computers in this sort of problem.

4.1 Tradeoff

There may well be a tradeoff between the amount of communications done now to improve the balance of the resulting embedding and the goodness of that embedding. The tradeoff would be expressed by the weighting the two costs; this weighting would depend on:

- The expected time to completion of the computation: the longer the computation is expected to run, the more we are willing to pay now for savings later.

- The volatility of the performance variations: the more variation to be expected, the less we would spend on making improvements in the embedding (since such improvements would be expected to be short lived). If other information is known about the variation in performance (for instance, that processors which slow down a little are expected to slow down more in the near future), this would also be useful.

- The ratio of costs involved in move a vertex of the computation graph versus the cost of communicating information between processors: the more expensive moving a vertex is, the less we would be willing to do it.

In what follows, however, we all ignore that, and assume that we would just like to maximally improve the embedding, though at a minimal cost.

27-8

4.2 Other Approaches

This problem is not as well studied as the static balance problem, but there has been some work done. For example, Song [22] has a an algorithm for asynchronous load balancing, though it assumes (in our teminology) a computation graph with no edges. Ferraro, Liewer and Decyk [4] and Williams [25] also address dynamic load balancing.

4.3 Minimum Cost Flow Approach

First, we run the computation for some time period with the initial embedding. Measuring the time each processor takes to run its part of the computation graph (and, potentially, how long each individual vertex of the computation graph takes) gives us a relative performance figure p_i for each processor, measured in (for example) seconds per vertex.

We can now use this information to compute how many vertices should be on each processor: each processor gets a number of vertices proportional to the number of vertices it processed in each time step (that is, the inverse of the speed). If this could be done exactly, and the speed of the processors did not change, then all processors would finish in exactly the same amount of time. Unfortunately, we must round to an integer number of processors, and the processors' effective speeds are likely to change.

The information about how many vertices should be on each processor and the current number of vertices on that processor, give us either a number of vertices which should be sent to that processor (a demand) or a number of nodes which that processor would like to get rid of (a supply) or neither. We can arrange rounding to be done so that total supply equals total demand, though this is not really necessary.

A minimum cost flow algorithm is then used to generate a flow which matches up the supplies and demands by paths in the processor graph. To realize this flow, a processor which is sending vertices pairs with one receiving vertices, and a variant of Kernigan-Lin is used to transfer the relevant number of vertices and reoptimize the partition between those two processors. (Using the Hendrickson and Leland [11] multiway partitioning would be better, but it is not obvious how to parallelize it.)

5 Implementation

The current implementation is far from optimal, as it was intended mostly as a proof-of-concept type system. It should also be pointed out that it does not actually do a real computation. The implementation runs on an Intel iPSC/860 multiprocessor, a 16-node hypercube at Phillips. The hypercube is connected (*via* a front-end) to an ethernet which also includes some Sun Sparcstations.

5.1 The Distributed Graph Data Structure

Each vertex is given a unique 32-bit identifier. Each processor keeps track of which vertices of the computation graph C it "owns", and those vertices adjacent to those vertices, along with which processor that vertex belongs to. Each processor keeps track of its vertices with an AVL-tree.

AVL trees are one of the many self-balancing trees around. They were chosen because of the existence of a freely usable implementation [2]. The relevant details about AVL trees is that a given node (by id number) can be found "quickly" (in time proportional to the $\log_2$ of the number of elements in the tree), and the elements of the tree can be iterated over in linear time (by recursion).[6]

5.2 Simulated Load Variation

Load variation was done entirely in terms of simulated load on the processor; we did not do any variation in the cost of the computation at the vertices (though that would actually be more relevant to PIC codes, where

---

[6] Almost certainly there are better data structures for the job. However, I didn't have time to code anything myself; failing the existence of [2], I would have probably used something even worse.

the particles move between vertices of the communications graph). There were four different schemes used, all involving choosing a random "load" for the processor, then computing how long a simulated computation would take based on that time.

This information was sent to the front end for the min cost flow computation.

5.3 The Min Cost Flow Algorithm

The Min Cost Flow algorithm used based on LEDA, the Library of Efficient Discrete Algorithms [18]. This is a C++ library which includes a graph data structure[7] and some graph algorithms. A minimum cost flow algorithm is not included, but using the algorithms LEDA does include, we coded a minimum cost flow algorithm from Hu [12, page 169], specifically the one due to Klein [15]. This required the negative cost cycle finding algorithm in Hu [12, page 157].

Since there was no C++ compiler on the iPSC (either for the processors or the front end), we ran the Min Cost Flow algorithm on a Sun Sparcstation 10, and used the RPC facility[8] to communicate between the front end and the Sun.[9]

5.4 Results

Timing runs are shown in tables 1–4. The first half of the table shows the times without load balancing, the second with load balancing. The columns marked *min* and *max* show the times of the fastest and slowest processor, while the columns marked Δ show the difference in times. Finally, the *real time* columns show the real time required for the run (in seconds).

The graphs we used are:

grid4x4 A 4x4 square grid

grid10x10 A 10x10 square grid.

grid10x10x10 A 10x10x10 cubic grid.

grid20x20x20 A 20x20x20 cubic grid.

small A small "cactus", similar to 'test', but with a 3x3x6 "box" and 1x1 "arms".

test A larger cactus; see figure 3.

In each run, the graph was started with a partition generated by **Chaco**, using the multilevel method. In table 1, the ith processor was given a speed of $(16 - i)/4$ vertices per time step. In table 2, the same speeds were used, but with small random variations each time. (These two cases were run mostly to debug the programs, as it is fairly clear what the results should be.)

In table 3 and table 4, each processor maintained a simulated job count. The speed was generated by giving the simulated process 1/(number of jobs) time. In each case, the number of jobs changes randomly each time step; in the second, the speed was further modified by a small amount of random noise.

6 Conclusions

Based on our results, it seems that dynamic load balancing can give large speed ups in computationally feasible amounts of time. More work is needed.

---

[7]That's two.

[8]And a third graph representation

[9]Needless to say, no claim that this is the most efficient scheme is being made! It obviously would have been better to avoid the front end, but the nodes did not have RPC facilities available.
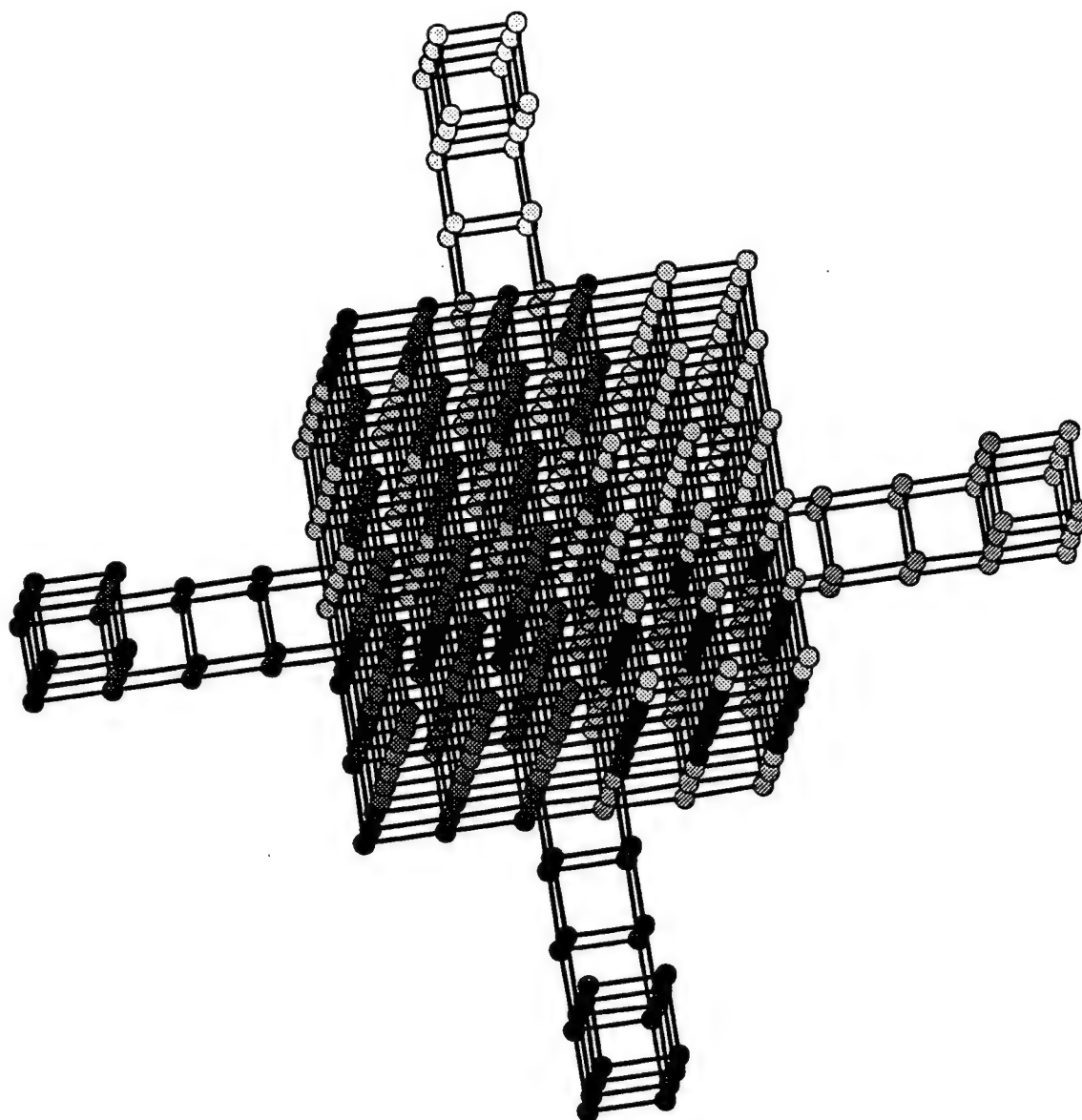
Figure 3: The cactus graph "test". This graph (and the others used in this paper) was generated by a small grid generator written by the author, which reads a 'little language' for describing cubic grids and outputs a graph representing that grid in **Chaco** input format. (The colors represent a partition of the cactus generated by **Chaco**.) The figure was drawn with the help of a graph editing utility also written by the author. (Both of these were written while at Phillips.)

file	-lnone				-lflow			
	min	max	Δ	real time	min	max	Δ	real time
grid10x10	35	480	445	5.980	120	288	168	26.555
grid10x10x10	315	5040	4725	43.429	1243	2224	981	82.379
grid20x20x20	2500	40000	37500	290.290	9960	17536	7576	577.349
grid4x4	0	320	320	2.243	0	320	320	11.640
small	145	2240	2095	20.342	565	1035	470	47.882
test	2540	40640	38100	282.223	10116	17792	7676	572.364

Table 1: Results for static processor speeds

file	-lnone				-lflow			
	min	max	Δ	real time	min	max	Δ	real time
grid10x10	88	309	221	6.952	169	198	29	69.372
grid10x10x10	791	3247	2456	44.439	1668	1898	230	162.392
grid20x20x20	6278	25773	19495	290.647	13377	15224	1847	712.788
grid4x4	0	206	206	3.086	0	201	201	32.508
small	364	1443	1079	20.899	761	859	98	122.571
test	6378	26185	19807	283.353	13586	15472	1885	704.784

Table 2: Results for static processor speeds with random variation

file	-lnone				-lflow			
	min	max	Δ	real time	min	max	Δ	real time
grid10x10	330	686	356	9.257	477	580	103	62.907
grid10x10x10	3410	6944	3534	46.024	4838	5244	406	155.676
grid20x20x20	27500	56000	28500	293.060	38774	42087	3313	593.562
grid4x4	0	448	448	4.692	0	356	356	34.831
small	1540	3248	1708	22.625	2203	2421	218	110.759
test	27885	56896	29011	287.461	39362	42741	3379	586.691

Table 3: Results for dynamic proc speeds

file	-lnone				-lflow			
	min	max	Δ	real time	min	max	Δ	real time
grid10x10	88	309	221	6.953	169	198	29	69.314
grid10x10x10	791	3247	2456	44.279	1668	1898	230	162.393
grid4x4	0	206	206	3.107	0	201	201	32.473
small	364	1443	1079	20.965	761	859	98	136.427
test	6378	26185	19807	284.090	13586	15472	1885	710.428

Table 4: Results for dynamic proc speeds with random variation. (In this case, we were not able to run the grid20x20x20 graph to to lack of time.)

7 Future Research

Many areas of future research suggest themselves. For instance:

- Instead of using a cubic grid, the actual PIC graph could be used.

- Instead of (or in addition to) simulating processor load, the movement of particles could be simulated to give varying amount of computation to be done on each vertex.

- The algorithm could be tried on real PIC problems, and on different multiprocessors. Those with different communications architectures would be especially interesting.

- The min cost flow algorithm used is not the best known (see, for instance, Goldberg [7]). Furthermore, it ought to be done on the parallel processor. This suggests investigating parallel min cost flow algorithms; in particular, such algorithms where there is one processor for each node in the graph and communications links and edges also correspond.

- It is not clear that the min cost flow algorithm is needed at all. A local algorithm like that of Song [22] (modified to take account of communications requirements) might produce similar speed ups at a lower cost.

- Parallelizing Hendrickson and Leland [11]'s n-way Kernigan-Lin-type algorithm would be useful.

8 Acknowledgements

References

[1] STEPHEN T. BARNARD and HORST D. SIMON, Fast multispectral implementation of recursive spectral bisection for partitioning unstructured problems, *Concurrency*, 6(2):101–117, April 1994.

[2] KENDALL BENNETT, C tools, from The Net.

[3] C. K. BIRDSALL and A. B. LANGDON, *Plasma Physics via Computer Simulation*, Adam Hilger, New York, 1991.

[4] ROBERT D. FERRARO, PAULETT C. LIEWER, and VIKTOR K. DECYK, Dynamic load balancing for a 2D concurrent plasma PIC code, *Journal of Computational Physics*, 109:329–341, 1993.

[5] C. M. FIDUCCIA and R. M. MATTHEYSES, A linear-time heuristic for improving network partitions, in *19th IEEE Design Automation Conference*, pages 175–181, 1982.

[6] M. R. GAREY and D. S. JOHNSON, *Computers and Intractability*, Freeman, New York, 1979.

[7] ANDREW V. GOLDBERG, An efficient implementation of a scaling minimum-cost flow algorithm, preprint, August 1992.

[8] BRUCE HENDRICKSON and ROBERT LELAND, An improved spectral graph partitioning algorithm for mapping parallel computations, Technical Report SAND92-1460, Sandia National Laboratories, Albuquerque, NM, September 1992.

[9] BRUCE HENDRICKSON and ROBERT LELAND, The Chaco user's guide version 1.0, Technical Report SAND93-2339, Sandia National Laboratories, Albuquerque, NM, October 1993.

[10] BRUCE HENDRICKSON and ROBERT LELAND, Multidimensional spectral load balancing, Technical Report SAND93-0074, Sandia National Laboratories, Albuquerque, NM, January 1993.

[11] BRUCE HENDRICKSON and ROBERT LELAND, A multilevel algorithm for partitioning graphs, Technical Report SAND93-1301, Sandia National Laboratories, Albuquerque, NM, October 1993, draft.

[12] T. C. HU, *Integer Programming and Network Flows*, Addison-Wesley, Reading, Massachusetts, 1969.

[13] B. W. KERNIGAN and S. LIN, An efficient heuristic procedure for partitioning graphs, *Bell Systems Technical Journal*, **49**:291–307, February 1970.

[14] S. KIRKPATRICK, C. D. GELATT, and M. P. VECCHI, Optimization by simulated annealing, *Science*, **220**(4598):671–680, May 1983.

[15] M. KLEIN, A primal method for minimal cost flows, *Man. Sci.*, **14**(3):205–220, November 1967.

[16] JOSEPH W. H. LIU, A graph partitioning algorithm by node seprators, *ACM Transactions on Mathematical Software*, **15**(3):198–219, September 1989.

[17] CAPT BRYAN M. MINOR, Proposed revisions to PICARD, preprint, May 1994.

[18] STEFAN NÄHER, Leda user manual version 3.0, comes with LEDA.

[19] ALEX POTHEN, HORST D. SIMON, and KANG-PU LIOU, Partioning sparse matrices with eigenvectors of graphs, *SIAM Journal of Matrix Analysis and Applications*, **11**(3):430–452, July 1990.

[20] JOHN E. SAVAGE and MARKUS G. WLOKA, Parallelism in graph-partitioning, *Journal of Parallel and Distributed Computing*, **13**:257–272, 1991.

[21] CAPT BRAD SMITH and LOUIS BAKER, Parallel smooth particle hydrocode, talk given at Phillips.

[22] JIANJIAN SONG, A partially asynchronous and iterative algorithm for distributed load balancing, *Parallel Computing*, **20**:853–868, 1994.

[23] PETER R. SUARIS and GERSHON KEDEM, An algorithm for quadrisection and its application to standard cell placement, *IEEE Transactions on Circuits and Systems*, **35**(3):294–303, March 1988.

[24] DAVID W. WALKER, Characterizing the parallel performance of a large-scale, particle-in-cell plasma simulation code, *Concurrency*, **2**(4):257–288, December 1990.

[25] ROY D. WILLIAMS, Performance of dynamic load balancing algorithms for unstructured mesh calculations, *Concurrency*, **3**(5):457–481, October 1991.

ON THE CONSTITUTIVE BEHAVIOR OF SOLID PROPELLANTS

Krishnaswamy Ravi-Chandar
Associate Professor
Department of Mechanical Engineering

University of Houston
Houston, TX 77204-4792

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

On the Constitutive Behavior of Solid Propellants

Krishnaswamy Ravi-Chandar
Associate Professor
Department of Mechanical Engineering
University of Houston

Abstract

The constitutive behavior of solid propellants was studied. The materials are traditionally considered to be viscoelastic, with either linear or nonlinear models being used. A number of such models exist in the literature. In the present work, experiments were performed on a composite propellant to examine the mechanical behavior in detail and to compare the validity of the various models. Uniaxial tensile tests were performed at strain rates varying from 2×10^{-3} to 10^3 per min; relaxation tests were performed with strain levels from 4 to 20%; finally load-relax-unload and load-unload experiments were performed over a wide range of strain rates, strain levels and relaxation times. This large data set was all obtained from one batch of solid propellant. The results are reported here in detail with some preliminary interpretations in terms of the models. A detailed examination of the various models is being pursued and will be reported elsewhere.

ON THE CONSTITUTIVE BEHAVIOR OF SOLID PROPELLANTS

K. Ravi-Chandar

Introduction

Solid propellants exhibit an interesting range of nonlinear, time and load path dependent behavior; this range of behavior is quite complex and apparently many of the models that exist do not satisfactorily predict all aspects of the propellant mechanical behavior. This appears to be particularly true in the regions near crack tips where the strains are large, and the material is heavily damaged. However, it is necessary to have such a predictive capability in order to be able to handle more difficult problems regarding failure, degradation etc. In the following, we describe the results of a number of experiments that were aimed at elucidating the mechanisms of deformation that lead to the various complex material response characteristics observed.

Nonlinear Viscoelastic Models

Solid propellants are generally regarded as viscoelastic materials; it is common practice to characterize viscoelasticity using a linear theory and to express it in terms of the relaxation modulus, as determined from a relaxation test performed at 3% strain. Standard procedures for this have been written into the JANNAF handbook. The limitations of the linear theory at larger times has been recognized and there are quite a few attempts towards developing a nonlinear viscoelastic theory. (see for example, Farris, 1971, Schapery, 1982). Nonlinear theories have also been developed by others (Bernstein et al, 1963, McKenna and Zapas, 1979, Knauss and Emri, 1981, Boyce et al, 1988, O'Dowd and Knauss, 1992) for application to other polymeric materials. One of the main ideas of nonlinear viscoelastic theories is the introduction of a material clock or shifted time, where the shift depends on the volume dilatation, the stress or the strain level etc. Here we shall review two nonlinear models and their reduction to linear models; then we will discuss the experiments and results.

A general one dimensional nonlinear constitutive law, based on a free volume model can be described as follows:

$$\sigma(t) = \int_{-\infty}^{t} E\{\xi(t) - \xi(\tau)\}\frac{\partial \varepsilon(\tau)}{\partial \tau} d\tau \tag{1}$$

$E(t)$ is the relaxation modulus, $\varepsilon(t)$ is the imposed strain rate and $\xi(t)$ is a shifted time based on a material clock concept and is given by

$$\xi(t) = \int_0^t \frac{ds}{\Phi\{s\}} \qquad (2)$$

Here, Φ is the shift factor that depends on the fractional free volume f according to the Doolittle equation:

$$\log \Phi = \frac{b}{2.303}\left(\frac{1}{f} - \frac{1}{f_0}\right) \qquad (3)$$

$$f = f_0 + \alpha dT + B\,\beta \qquad (4)$$

where f_0 is the initial fractional free volume, α is the coefficient of thermal expansion, B is a constant and β represents the source of the nonlinearity either in terms of the pressure or the volume dilatation depending on the model chosen; some additional generalizations include consideration of α as time dependent and replacing αdT with a convolution integral representation. The shift factor is generally obtained from experimental measurements of the relaxation modulus over a wide range of the parameter space.

Schapery (1982), has proposed another model, by considering a representation in terms of a nonlinear, elastic material and relating its behavior to the viscoelastic material through a pseudo strain that is defined through a superposition integral. The pseudo-strain is defined as

$$\varepsilon^e = \frac{1}{E_R} \int_0^t E(t - \tau)\frac{\partial \varepsilon(\tau)}{\partial \tau} d\tau \qquad (5)$$

where E_R is an arbitrary reference modulus and $E(t)$ is the relaxation modulus. It is then assumed that there exists a potential function $\phi(\varepsilon^e)$ in terms of the pseudo-strains such that the stresses are obtained from ϕ as:

$$\sigma(t) = \frac{\partial \phi(\varepsilon^e)}{\partial \varepsilon^e}. \qquad (6)$$

Schapery allows for ϕ to depend also on the amount of damage sustained by the material by using a Lebesgue norm of the pseudostrains. Typically, the relaxation modulus is assumed to be determined from standard 3% strain level relaxation test and interpreted in terms of a power law fit such as

$$E(t) = (E_1 + t^{-n})E_2. \qquad (7)$$

The underlying assumption appears to be that the relaxation behavior is completely characterized by the above and that the nonlinearities are introduced primarily through the representation in Eq (6). In this report, we present the results of a detained experimental investigation into one composite propellant and interpret the results in terms of these two models.

Experimental Methods

A number of experiments were performed under uniaxial tensile loading in an effort to determine the nonlinear, viscoelastic constitutive behavior of the composite propellant. The specimens were all 2.25 in long, 0.5 in wide and 0.2 in thick; aluminum end tabs were bonded with a two-part urethane adhesive resulting in a gage length of 2 in. The specimens were then subjected to different time histories of uniaxial loading conditions, using two different Instron testing machines - a standard machine capable of crosshead rates from 0.005 in/min to 20 in/min and a high rate testing machine capable of crosshead rates up to 2500 in/min. Thus, the propellant specimens were loaded at strain rates ranging from 0.0025 per min to 1250 per min, varying over six orders of magnitude. The time histories imposed included monotonic ramping to failure, load-relaxation, load-relax-reload, load-relax-unload and load-unload tests. The strain levels were varied in the range from 4% to 20% and the strain rates for each segment of the load-unload tests were varied over the range indicated above. The strain was computed from the crosshead position data and the stress was calculated from the load cell data. Both the crosshead position and the load cell output were monitored as a function of time by a computerized data acquisition system connected to the Instron and were stored in a file. We present the experimental results in this section under the following three different categories and will then follow this with a discussion in terms of linear and nonlinear viscoelastic theories.

(i)   Relaxation modulus and its nonlinearites based on strain induced free volume changes

(ii)   Strain rate dependence of the propellant and relationship to relaxation modulus.

(iii)   Unloading behavior and its time dependence.

However, in order to set the stage for a discussion of the uniaxial behavior, a stress strain curve obtained using a ramp loading at a strain rate of 0.05 in per min up to 20% strain is shown in Figure 1. Two curves are shown, the circle symbols corresponding to the first loading to 20% followed by unloading and the triangle symbols corresponding to the second loading unloading cycle. A straight line is drawn through the initial loading segment to indicate the linear behavior of the stress strain curve at low strain levels. Clearly, a nonlinear behavior is evident beyond a strain of about 11%; this is due to the debonding of the particles from the rubber matrix and is called dewetting. Also the unloading and reloading curves clearly indicate the stress-strain curve is path-dependent. The unloading curve is quite steep indicating a significantly stiffer response than during the initial loading! Upon subsequent reloading the specimen exhibits a smaller modulus; this effect is called "Mullin's effect" and has been studied extensively in lowly filled particulate rubbers. Note that the

second unloading is not affected by the previous load history and appears similar to the first unloading.

Relaxation Modulus

The relaxation modulus of solid propellants is typically determined in a dog-bone specimen as per the guidelines of the JANNAF Handbook on the Mechanical Behavior of Solid Propellants. This test requires that a strain of 3% be applied and held for a long time; while the rate of straining to 3% is not important, relaxation modulus is typically computed only a long time after the strain of 3% is reached. The relaxation modulus is then computed as the instantaneous stress divided by the constant strain. By determining the relaxation modulus at different temperatures and using the time-temperature superposition, a master relaxation modulus curve is typically generated that covers several decades of time. This master relaxation curve is then assumed to provide a complete picture of the relaxation behavior of the solid propellant at all times, temperatures and strain levels using appropriate shift factors based on temperature. Since a large collection of relaxation data at 3% strain is readily available for several of the composite propellants, and since they all appear to be representable over several decades of time by a simple power law of the type indicated in Eq.(7), our aim was to investigate the nonlinearities induced by strain. Thus, nonstandard relaxation experiments were performed at strain levels of .04, 0.08, 0.12, 0.16 and 0.20; the strain rate used in reaching this strain was 0.05 per min. The time range over which the relaxation data were obtained varied from about 120 min to about 800 min. The relaxation modulus variation with time from these measurements are shown in Figure 2. Clearly, the relaxation moduli at different strain levels exhibit very nearly the same time dependence, but are shifted in time. This indicates a dependence of the internal material clock on the strain; thus, the free volume in Eq.(4) can be taken to depend on the strain level as

$$f = f_0 + \alpha dT + B \ (\varepsilon) \tag{8}$$

where $B(\varepsilon)$ is to be determined from the experimental data in Figure 2 in order to provide a unique master relaxation curve. We shall represent the strain effect on the relaxation modulus directly through a determination of the dependence of the shift factor on the strain level; the latter is obtained merely by obtaining a smooth master curve from all the relaxation data at different strain levels. The resulting master relaxation curve and the variation of the shift factor Φ with strain level are shown in Figures 3 and 4 respectively. The master curve exhibits a power law type behavior and can be represented as

$$E = 724t^{-0.14} \tag{9}$$

The shift factor variation with strain, however, exhibits behavior not observed heretofore. A dashed curve is drawn through the data only to exhibit the trend and is not a precise fit of the data; it is necessary to repeat the measurements a number of times to determine statistical significance of each measurement. Most observations of strain induced shift factor indicate a monotonic variation with strain (Gondouin, 1993), while the present results indicate an abrupt change at round 12% strain. It is interesting to note that this strain level coincides with the level at which dewetting begins (see Figure 1), and beyond which the uniaxial stress strain curve is nonlinear. Using a free volume approach, one might speculate that while at strains below 12%, the global strain might be directly related to the volume change, above 12% strain vacuole generation and growth cause bulk of the volume change and that the actual free volume change in the binder is not easily related to the global strain. More investigations are necessary to explore these ideas and characterize the nonlinear relaxation behavior of the solid propellant.

Strain Rate Dependence of the Modulus

The suitability of the characterization of the material relaxation by Eq.(9) can be determined easily by determining the uniaxial stress-strain behavior at different strain rates using Eqs.(1) and (9) and comparing them with the measured stress-strain curves from uniaxial tests at different rates of loading. In performing this comparison, we shall restrict attention initially to the linear portion of the stress-strain behavior. In Figure 5, the experimentally measured stress strain curve from six different tests at strain rates ranging from 0.0025 per min to 1250 per min are shown. Clearly, each of these curves exhibits a linear region, but the nonlinearity begins at smaller strains when the strain rate is increased. The initial modulus of the material was determined (by linear regression) from the slope of the linear portion for each of the fifteen tests. This initial modulus is plotted as a function of time obtained as the inverse of the strain rate and is shown in Figure 6; the initial modulus was also calculated by computing the stress strain curve from the viscoelastic constitutive law in Eq. (1), using Eq.(9) for the relaxation modulus and is also shown in Figure 6. There appears to be a large scatter in the high strain rate (short time) experiments, but clearly the trend is that the relaxation behavior extracted from long time data cannot be extended easily into the short time range; however, note that while the relaxation tests were only performed over three decades of time, the high strain rate tests indicate that the simple power law can be extended into a larger range of at least five decades and perhaps more into the long time region. In the short time region, additional experiments are needed, perhaps at different temperatures, so as to determine the complete relaxation modulus; however, this time scale of <0,1 min is only of interest under limited special circumstances. The nonlinear portion of the stress-strain curve in Figure 5 and the load-relaxation test results are still under investigation to determine the applicability of the constitutive behavior in Eqs. (1-4) or Eqs. (5-9).

Unloading Behavior

As indicated in Figure 1, the unloading occurs rapidly: the load drops dramatically for a small drop in the strain. A linear viscoelastic analysis would not predict such a behavior; since this unloading behavior is also evident at small strains where linear viscoelasticity is appropriate, one must look at some other mechanism that contributes to this rapid unloading. To explore this further, repeat experiments were performed, where the specimen was loaded to 20% strain at a rate of 0.05 per min and then unloaded at various rates ranging from 0.005 to 5 per min. Figure 7 shows a comparison of the results from these tests; the unloading rates in strain per min is given in the legend. The loading part of the stress strain curve from all four tests fall on each other indicating the degree of repeatability of these experiments. Also, the unloading behavior appears to be unaffected by time scale of the unloading - even when the specimen is unloaded at rates differing by four order of magnitude; this cannot be true even of linear viscoelasticity governs the behavior of the material. It was indicated earlier that the unloading behavior also did not depend on the nature of the previous loading (see Figure 1). Thus the mechanism of the deformation during unloading must be substantially different from the mechanism during the loading or relaxation; one might speculate now that during unloading, the ligaments of rubber that were stretched to large strains during the loading phase, snap back elastically during a very small change in the global strain (the local strains in the rubber could change by a much larger amount). This speculation also needs to be further investigated in order to determine the complete deformation mechanism during unloading. These results and some additional experiments at various loading histories not discussed here are being examined further in terms of the constitutive models in Eqs.(1-9) and will be reported elsewhere.

Conclusions

The constitutive behavior of a composite solid propellant was examined by performing uniaxial tests under various strain histories. Three main conclusions are at hand and further investigations using these data are in progress.

(i)   The relaxation behavior of the solid propellant exhibits a strain-induced nonlinearity that is not easily interpreted in terms of the present models.

(ii) The strain rate dependence of the modulus was shown to be representable by the relaxation modulus in the long time range, but not when the rates of loading became very large.

(iii) The unloading stress-strain behavior was shown to be independent of the previous load history and this observation poses interesting challenges to the viscoelastic models.

Acknowledgment

References

Bernstein, B., Kearsley, E.A. and Zapas, L.J., 1963, "A study of stress relaxation with finite strains", *Transactions of the Society of Rheology*, **7**, pp.391-410.

Boyce , M.C., Parks, D.M., and Argon, A.S., 1988, "Large inelastic deformation of glassy polymers: Part I: Rate dependent constitutive model", *Mechanics of Materials*, **7**, pp. 15-33.

N.P. O'Dowd and Knauss, W.G., "Time dependent large principal deformation of polymers", GALCIT SM-92-40, California Institute of Technology, Pasadena.

Farris, R.J., 1971, "The stress-strain behavior of mechanically degradable polymers", in **Polymer Networks: Structural and Mechanical Properties**, ed. A.J. Chompff and S. Newman, Plenum Publishing Corporation, pp.341-394.

Goundouin, B., 1993, "Structural analysis of propellant grains", in **Solid Rocket Propulsion Technology**, edited by A.Davenas, Pergamon Press, Oxford, pp.215-302.

Knauss, W.G., and Emri, I., 1981, "Nonlinear viscoelasticity based on free volume considerations", Computers and Structures, **13**, pp.123-128.

McKenna, G.B., and Zapas, L.J., 1979, "Nonlinar viscoelastic behavior of poly(methyl methacrylate) in torsion", Journal of Rheology, **23**, pp.151-166.

Schapery, R.A., 1982, "Models for damage growth and fracture in nonlinear viscoelastic particulate composites", Proceedings of the Ninth U.S. National Congress of Applied Mechanics, pp. 237-245.
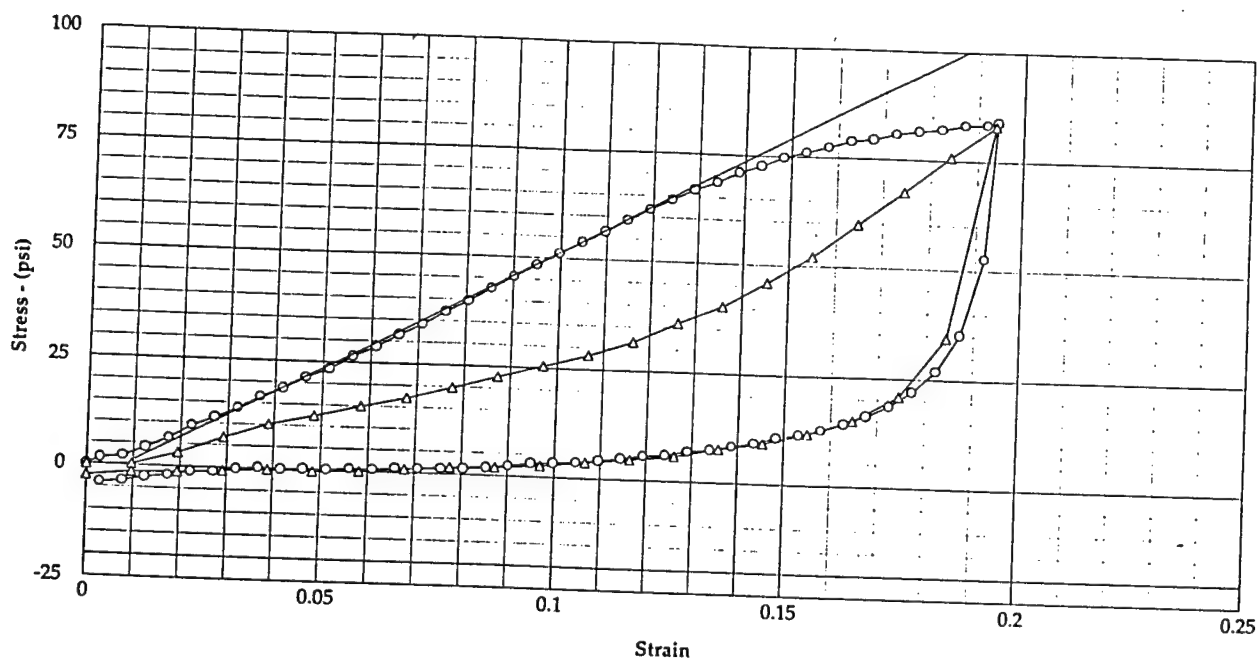
Figure 1. Uniaxial stress strain curve for a composite solid propellant. Two loading-unloading cycles are show. The circles indicate the first cycle and the trianges represent the second cycle. The rate of loading was .05 per min.
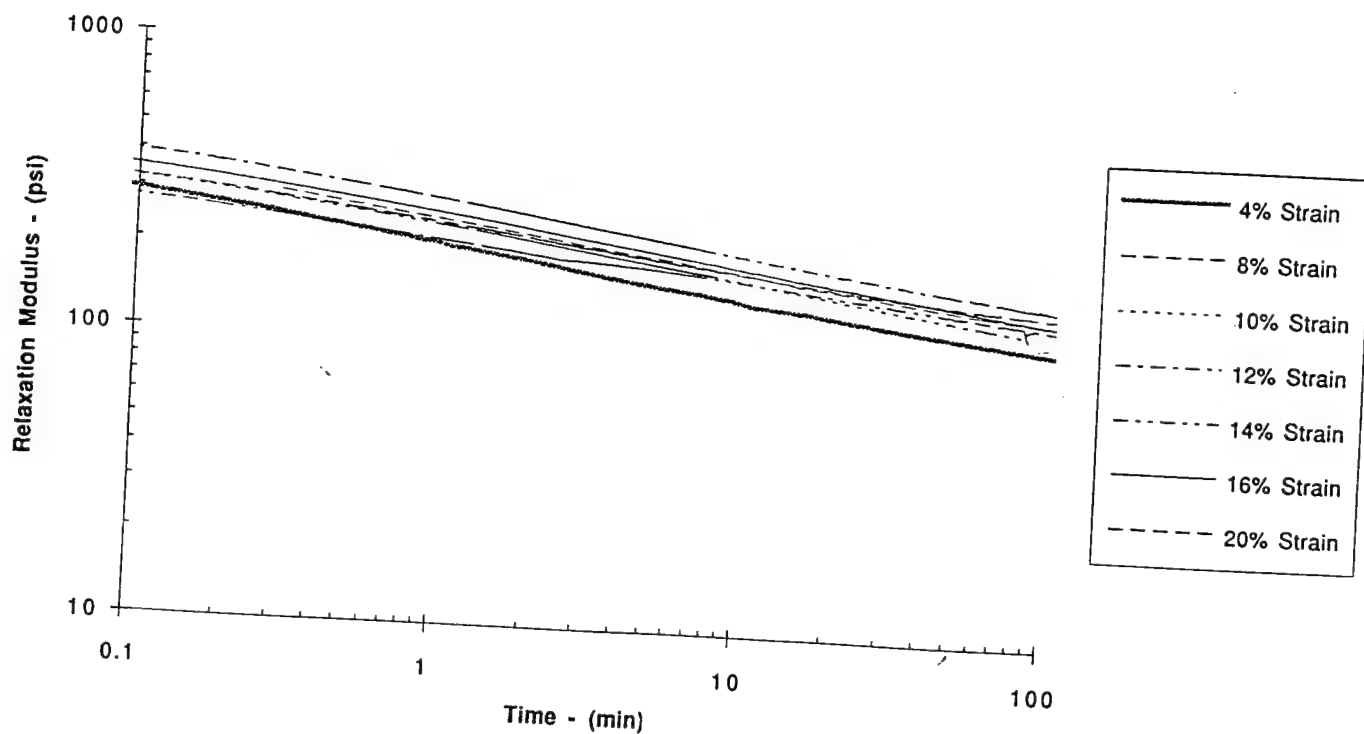
Figure 2. Relaxation modulus variation with time for different strain levels.
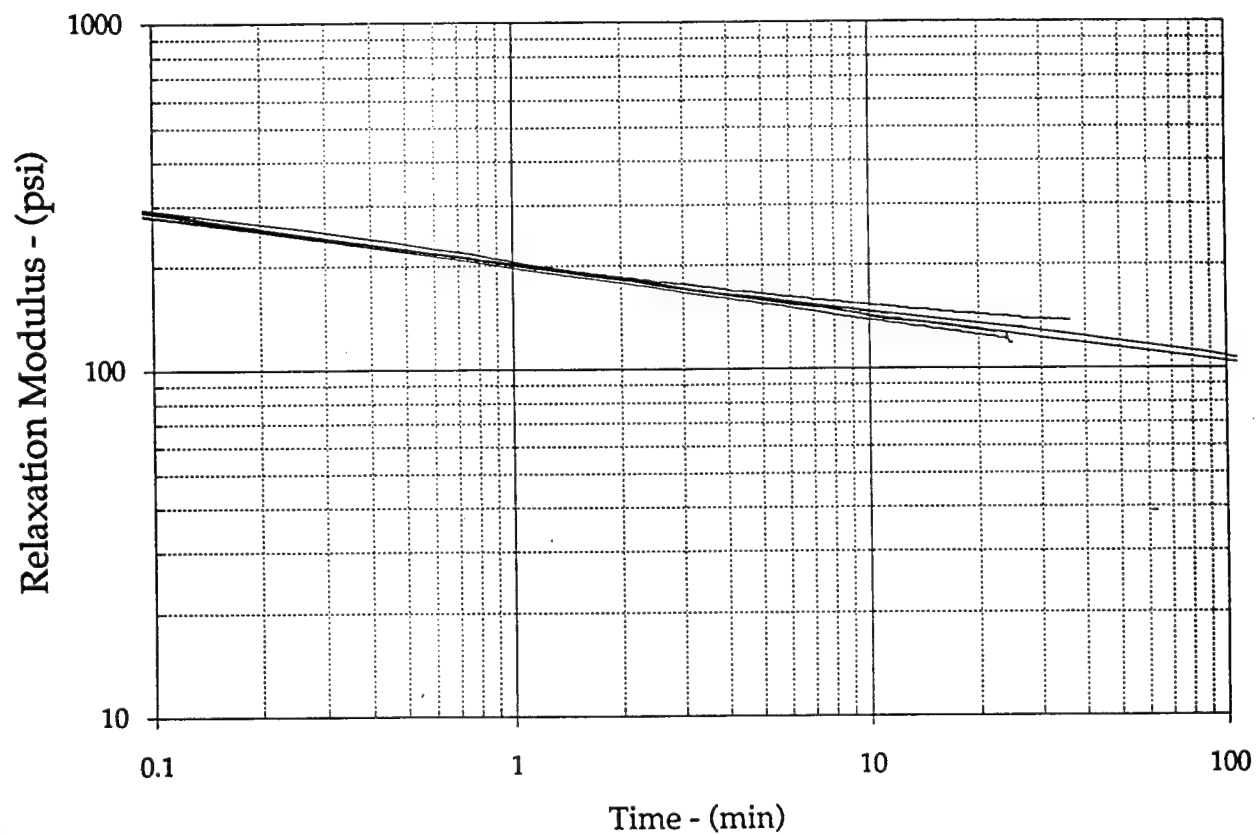
Figure 3. Master relaxation curve for the propellant, obtained by shifting the data in Figure 2 using a strain based shift factor shown in Figure 4.
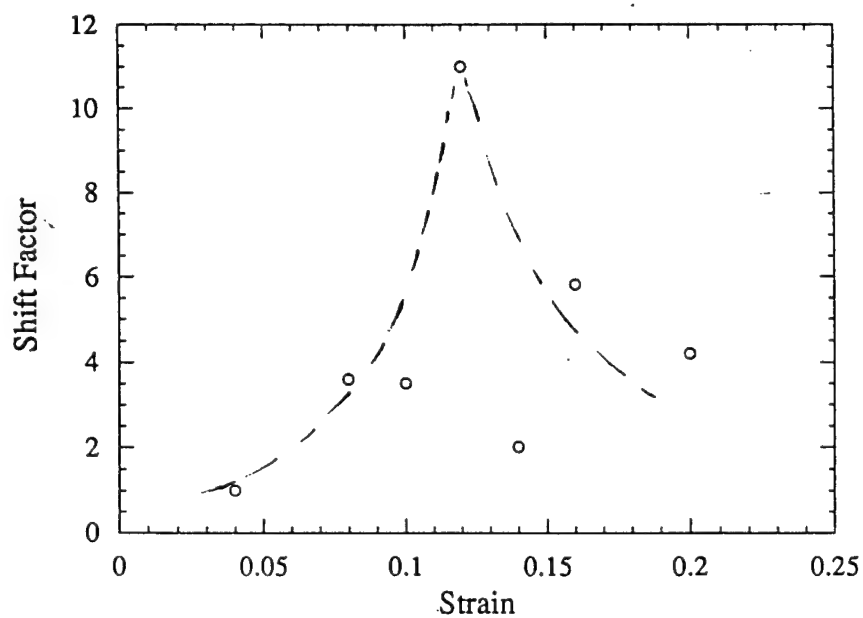
Figure 4. Dependence of the shift factor on strain; note that dashed line does not represent a quantitative fit, but a qualitative guideline for the trend in the data.
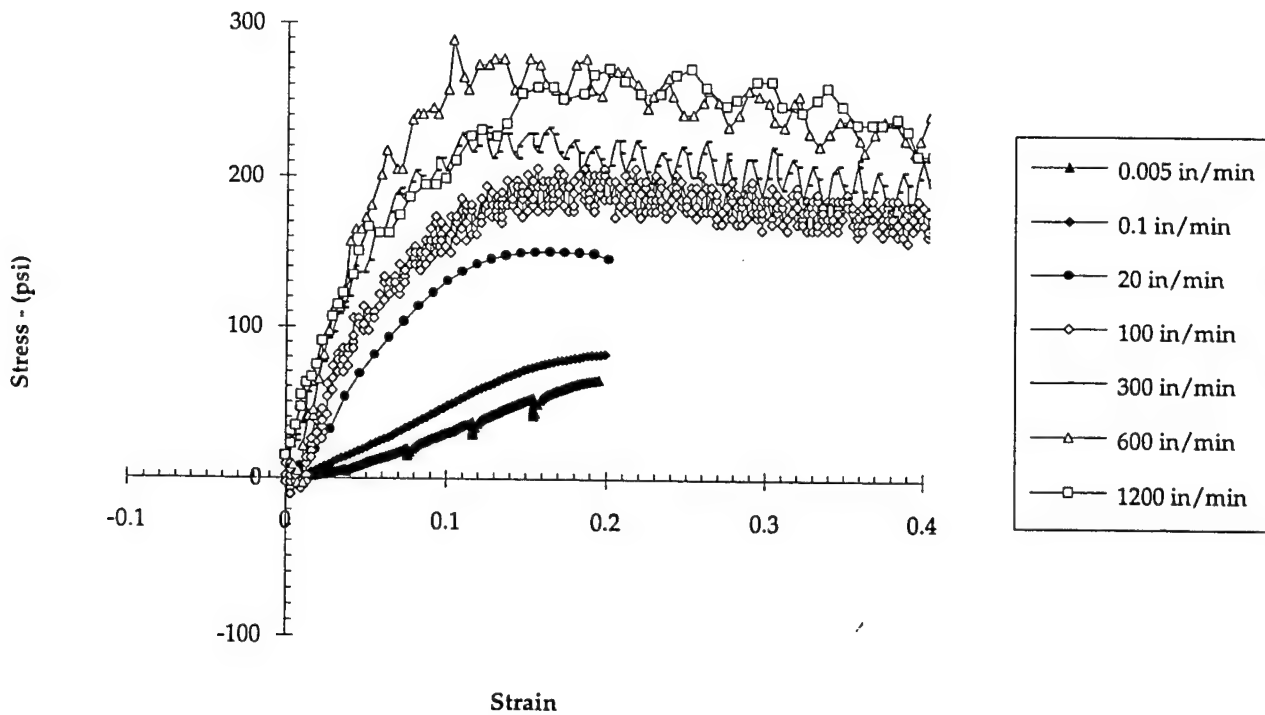
Figure 5. Dependence of the stress-strain curve on the strain rate. (Note that the legend indicates the cross head rate in in/min; the strain rate is one half of this value since the gage length was 2 in)
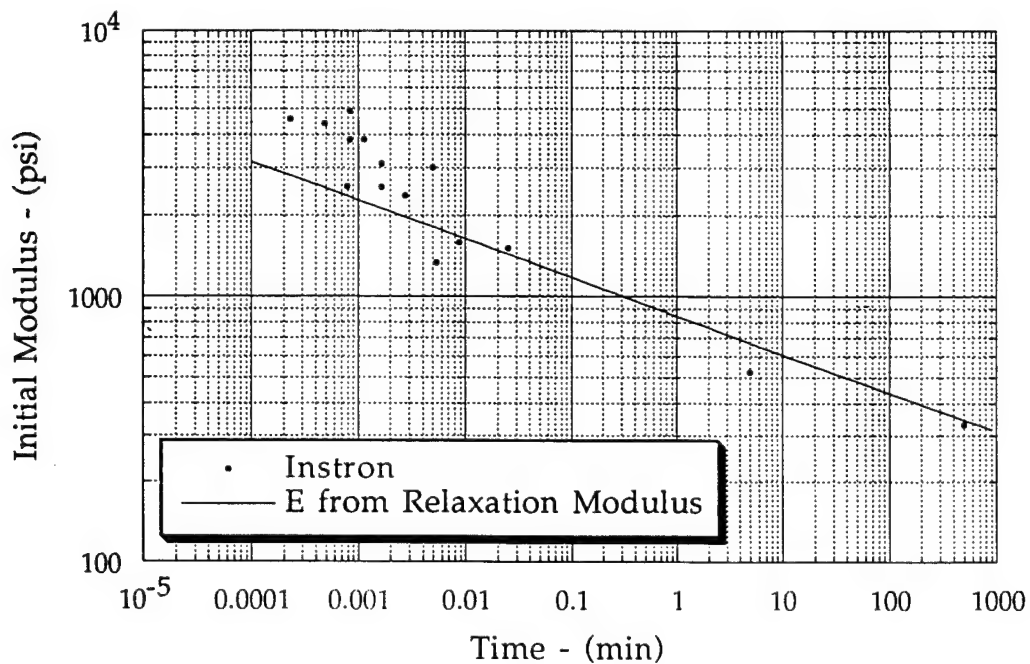
Figure 6. Dependence of the modulus on the strain rate obtained from high rate tests and a comparison with the predictions from the relaxation modulus.

28-12
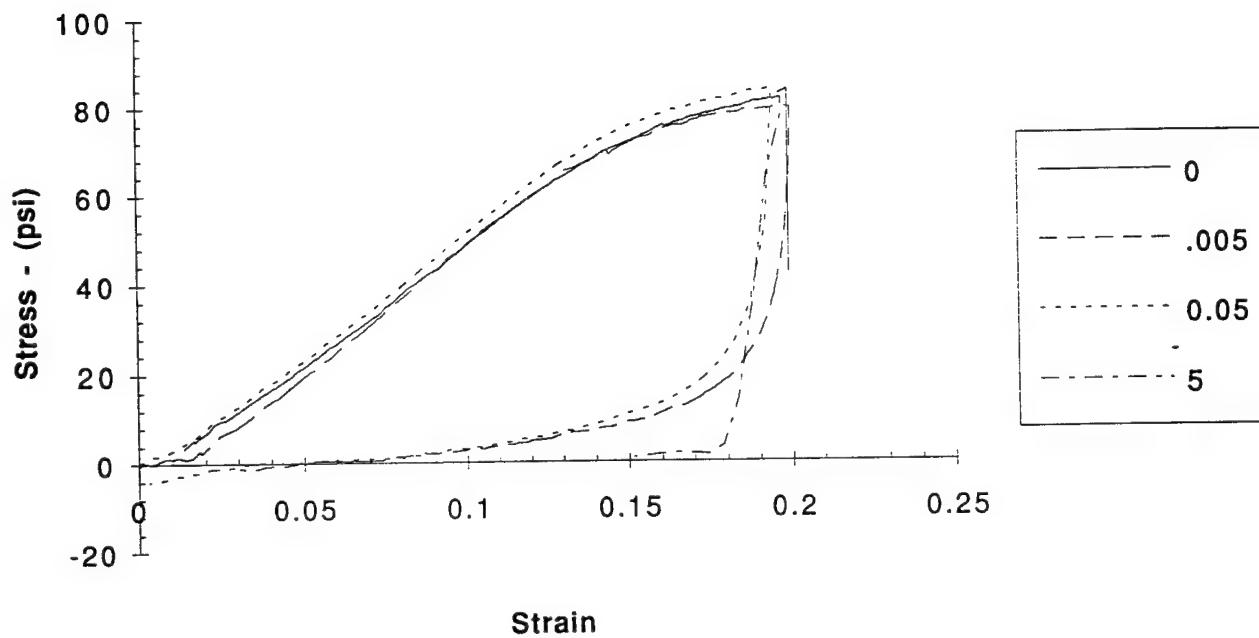
Figure 7. Unloading stress-strain behavior and its independence on the rate of unloading.

RELAXATION PROCESSES IN GAIN SWITCHED IODINE LASERS

Jeff Nicholson
Graduate Student
Wolfgang Rudolph
Associate Professor
Department of Physics and Astronomy
The University of New Mexico
Albuquerque, NM 87131

29 − 1

RELAXATION PROCESSES IN GAIN SWITCHED IODINE LASERS

Jeff Nicholson
Graduate Student
Wolfgang Rudolph
Associate Professor
Department of Physics and Astronomy
The University of New Mexico
Albuquerque, NM 87131

Abstract

The dynamics of a gain switched, low pressure photolytic iodine laser were investigated experimentally and theoretically. The pulse shape, build-up and decay time were measured as function of the pressure of the active medium (CF_3I) and of the pressure of the buffer gas. At low pressure, the pulse develops a second peak, and the build-up and decay times become longer. The distinct features of the iodine pulse could be explained in terms of two relaxation processes involving collisions that change the direction and magnitude of the velocity vector. A computer model based on the rate equations which includes collisional relaxations allowed us to derive quantitative estimates for the corresponding relaxation times as a function of pressure and buffer gas.

RELAXATION PROCESSES IN GAIN SWITCHED IODINE LASERS

Jeff Nicholson and Wolfgang Rudolph

1 Introduction

A great deal of interest has been devoted to the development of iodine lasers for different types of applications, see e.g., [1],[2],[3]. The potentials of chemical pumping are particularly attractive for applications requiring high-power. Normally, lasing is achieved between the hyperfine levels 4 and 3 of excited atomic iodine, I^*, which exhibits a total of six hyperfine transitions. The wavelength of this transition is 1.315 μm. The dominance of the 3-4 transition is because its magnetic dipole moment is about two times greater than that of the next largest transition (2-2). For the past twenty years a number of fundamental studies of iodine lasers have been undertaken involving longitudinal mode-beating and mode-locking [4],[5],[6]. Photolytic pumping employing an excimer laser is a convenient technique for laboratory operation of iodine lasers. However little effort has been devoted to the primary processes in photodissociation lasers working at low pressure. From Doppler spectroscopy experiments [7] and the theory of collisional relaxation [8] it is known that a short dissociation pulse creates a nonthermal velocity distribution of excited iodine. Depending on the parent molecule and buffer gas (if any) two different relaxation types with different time constants can be distinguished - collisions that change only the direction of the velocity vector but not its magnitude with a time constant T_b and those determining the rate by which the velocity relaxes towards its equilibrium value (time constant T_v). A photolytic, single mode iodine laser is thus expected to experience a dramatic change in gain during the evolution of the gain switched pulse. Since in a single mode laser, saturation of the gain is associated with hole burning, collisional cross relaxation will effect the shape and duration of the laser pulse.

The aim of this project was to study the evolution of gain switched pulses in low - pressure photolytic iodine lasers with and without buffer gas. From a comparison of theory and experiment the possibility of determining relaxation constants was explored.

2 Experimental results

2.1 Experimental setup

The experimental setup is shown in Fig. 1. The active medium used in this experiment was CF_3I. The length of the gain cell was 85 cm. The cavity consisted of a mirror with a 5 m focal length and a mirror with a 25 m focal length, both with 100% reflectivity. With a cavity length of 1.1 m, this gave a radius for the beam waist of 1.1 mm and a confocal parameter of 2.9 m.
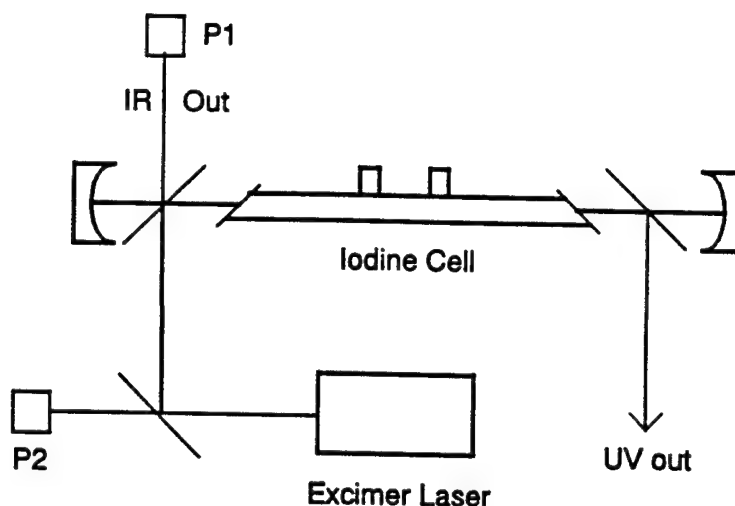
Figure 1: Experimental setup of a photolytic iodine laser pumped by an excimer laser. P1 was a fast photodiode used to measure the iodine pulse profile. P2 was a photodiode used to trigger the oscilloscope.

The pump was a KrF excimer laser operating at 248 nm, with pulse widths of about 20 ns and maximum pulse energies of approximately 200 mJ. The pump was focused with a 2 m fused silica lens, and coupled into the cavity with a dichroic beam splitter. The excimer pulse was coupled out of the cavity with a second dichroic mirror, in order to avoid damaging the end mirror of the iodine cell when working in the low absorption regime (low CF_3I pressure). The second dichroic mirror was placed at Brewster's angle for 1.3 μm in order to reduce the cavity losses.

The iodine laser output was taken from the dichroic beam splitter that coupled the excimer pump laser into the iodine cavity. With the cavity arranged as outlined above, the cavity losses per round trip were approximately 6%, and the cavity lifetime was about 60 ns. The iodine laser, under these conditions, lased at pressures as low as 0.2 Torr.

Longitudinal mode beating is an inherent property of gain switched and Q switched iodine lasers operating at low pressures [5]; therefore, an aperture was used in the iodine cavity to control the number of longitudinal modes in the cavity. When the aperture was sufficiently small, ($\sim$1.5mm in diameter) not only was the iodine laser forced to run in TEM_{00} mode, but strong longitudinal mode beating was eliminated in approximately 50% of the shots.

After laser action, only about 90% of the molecular fragments recombine to form CF_3I, the rest of the fragments going into processes such as dimerization [2]. I_2 is a strong quencher of the $^2P_{1/2}$ state, leading to a degradation of the iodine laser output after multiple shots of the excimer laser. Therefore, the iodine cell was refilled after every fourth or fifth excimer pulse. Note, that while these processes proved to be important for multiple shot experiments, they do not play a significant role in single shot pulse dynamics because of their slow (e. g. $\sim 4 * 10^{-31} cm^6 mol^{-2} s^{-1}$ for dimerization) rate.

The iodine pulse was detected with a fast (rise time < 500 ps) photodiode and viewed on a digital oscilloscope with a 200 MHz analog bandwidth and a 1 Gs sampling rate. The oscilloscope was triggered with a signal from a second photodiode that was sensitive to the UV excimer pulse. The delay was then set digitally with the oscilloscope itself.

The pulse dynamics were measured as a function of pressure of CF_3I plus buffer gas (if any). The quantities of interest were:

- τ_1 The build up time of the iodine pulse. This was defined as the time between the excimer pulse peak and the time when the iodine pulse builds to 10% of its peak intensity.

- τ_2 The exponential rise time of the iodine pulse.

- τ_3 The duration of the iodine pulse. This was defined as the time from τ_1 to the point when the iodine pulse fell to 10% of its peak value. The iodine pulse did not decay in an exponential fashion at low pressures, making an exponential fit impossible.

2.2 Results without buffer gas

Initially, the iodine pulse dynamics were measured as a function of the pres·␣·␣e of CF_3I, without the use of any buffer gas. Although the iodine laser would lase at pressures as low as 0.2 Torr, under these conditions the laser was near threshold, and consequently, the pulse characteristics varied greatly from shot to shot. Therefore, measurements were generally made for pressures of 0.4 Torr and greater.

Typical pulse shapes are shown in Fig. 2 for pressures of 0.6, 2, and 6 Torr of CF_3I. At low pressures, the iodine pulse shows a distinct second peak, caused by repumping due to the velocity relaxation. As the pressure increases, the peak becomes less distinct, becoming a shoulder, and finally disappears altogether at pressures above 3 Torr.

The buildup and decay time of the iodine pulse as a function of CF_3I pressure are shown in Fig. 3(a). Fig. 3(b). shows the exponential rise time of the iodine pulse as a function of pressure. The characteristic shape of these plots can be explained by the increased relaxation rate as the pressure increases as will be detailed in Section 3.3. Note, that by changing the CF_3I pressure, we also change the amount of absorbed pump pulse energy and thus, the overall gain.

2.3 Results with buffer gas

The same measurements outlined in the previous section were made using a buffer gas. The cell was filled with 0.6 Torr of CF_3I, and Xenon was added on top of this. Typical pulse profiles are shown in Fig. 4. Pulse buildup time and duration as a function of buffer gas pressure are shown in Fig. 5. The behavior of the build up time and pulse duration – and even the overall pulse
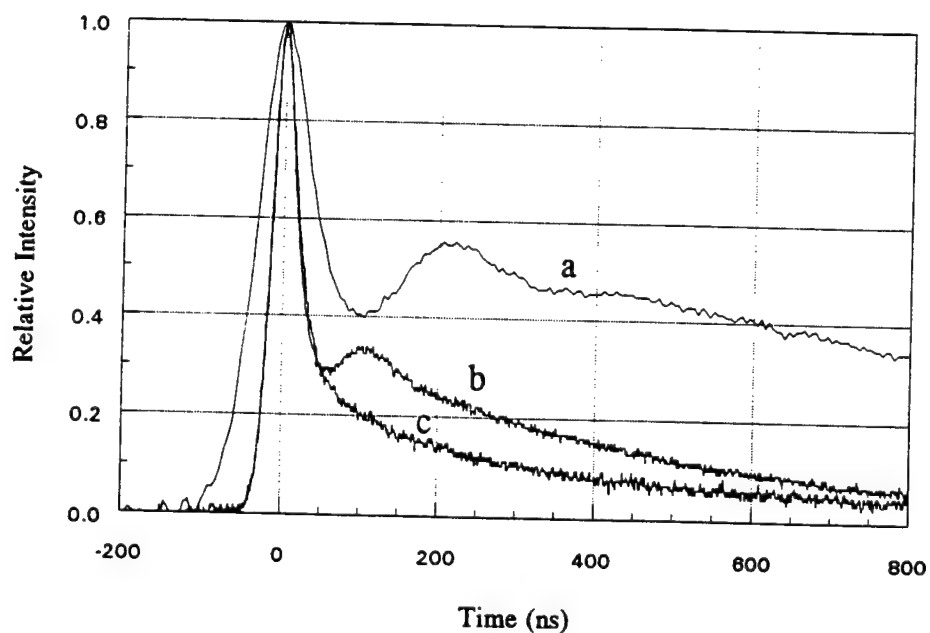
Figure 2: Pulse profiles for (a) 0.6 Torr, (b) 2 Torr, and (c) 6 Torr of CF_3I. The pulse peaks have been normalized to one and moved to the origin of the time axis for comparison purposes.

(a)

(b)
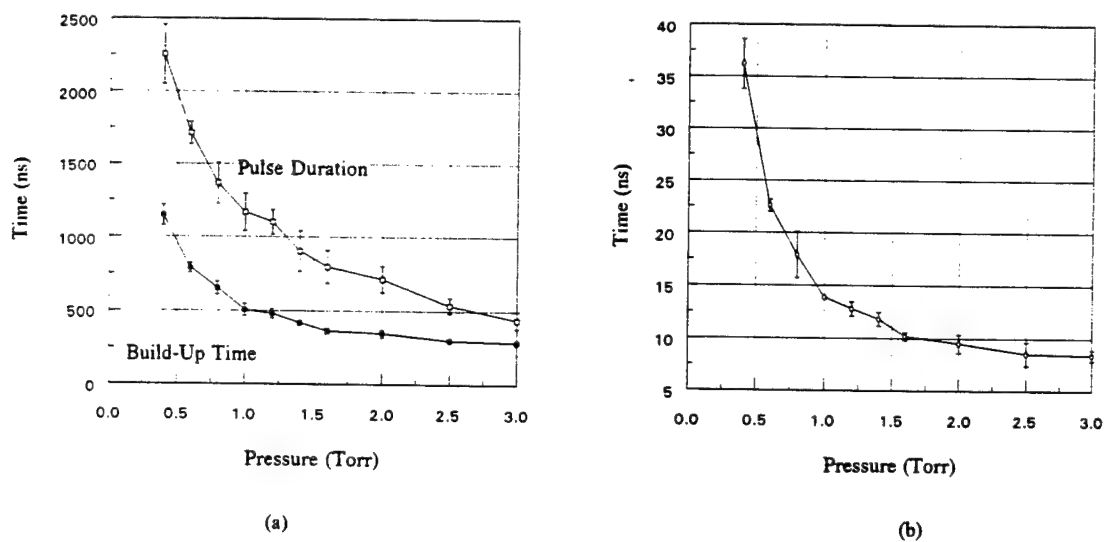
Figure 3: (a) Build up time and duration and (b) Rise time of the iodine pulse as a function of CF_3I pressure. No buffer gas was used.
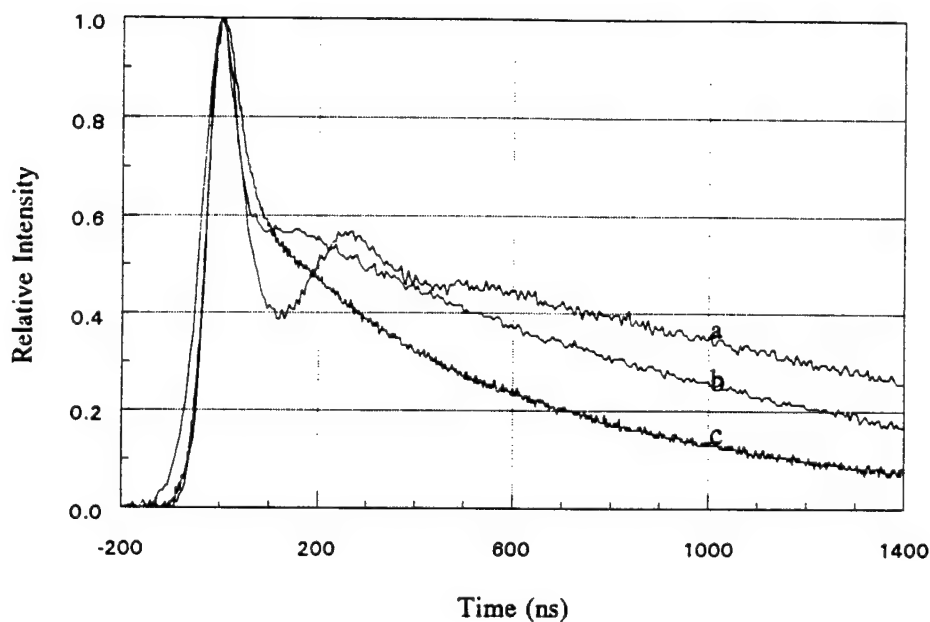
Figure 4: Pulse profiles for (a) 0.0 Torr, (b) 0.6 Torr, and (c) 1.2 Torr of Xenon added to 0.6 Torr of CF_3I. The pulse peaks have been normalized to one and moved to the origin of the time axis for comparison purposes.
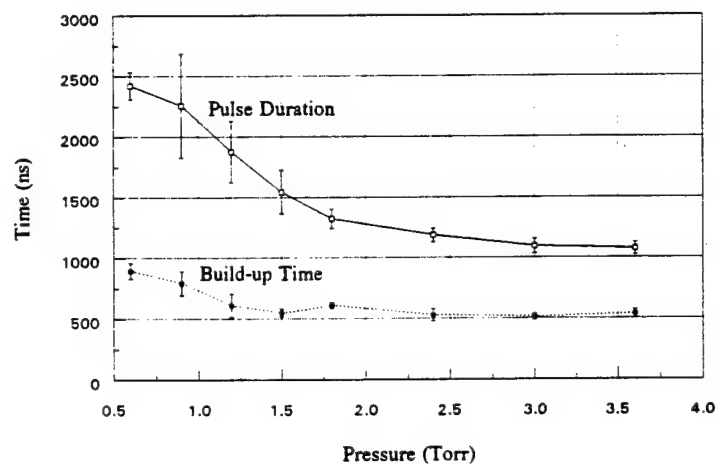
Figure 5: Build up time and duration of the iodine pulse as a function of buffer gas (Xe) pressure. CF_3I pressure was 0.6 Torr

29 – 7
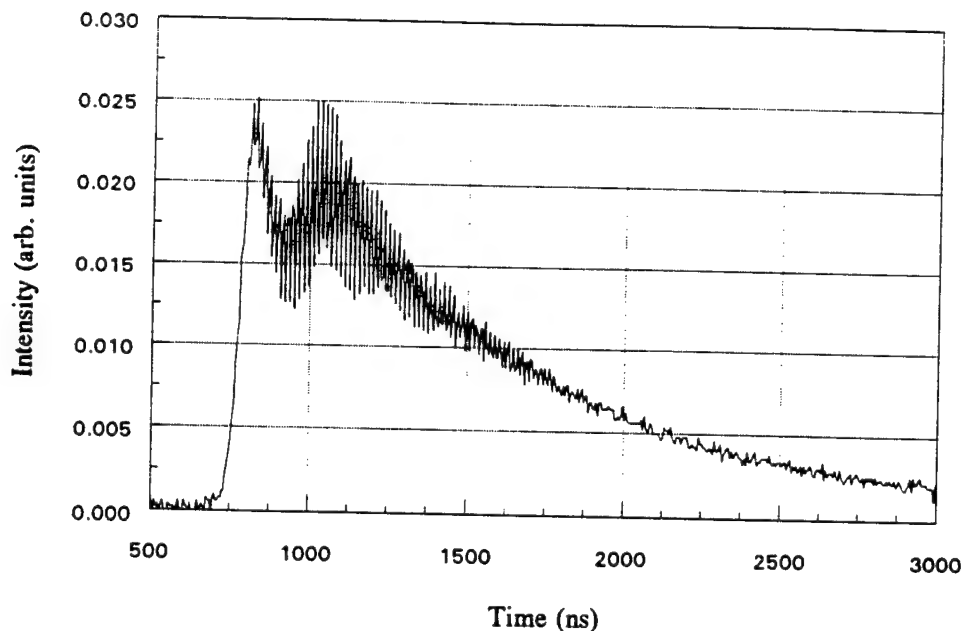
Figure 6: Typical iodine laser pulse showing two longitudinal mode exceeding threshold. CF$_3$I pressure was 0.6 Torr.

profiles – as a function of pressure is similar for pulses with and without buffer gas. However, the exponential rise time of the iodine pulse is essentially unaffected by the addition of Xenon. A discussion of these features and a comparison with the computer model are given in Section 3.3.

The effects of mode beating could never be fully eliminated in 100% of the shots. However, as mentioned above, the probability of mode beating could be decreased by reducing the aperture in the iodine cavity significantly, thus preventing all but one mode from exceeding threshold. A typical laser output in the regime where two modes can exist is shown in Fig. 6. The mode beating can clearly be observed, as can the fact that the second mode starts oscillating at a later time than the first. This is consistent with the higher losses this mode experiences. A complex beating of both longitudinal and transverse modes was observed when the iodine aperture was fully open.

3 Theoretical description of low pressure photolytic iodine lasers

3.1 Relaxation processes

A rate equation approach was chosen for the simulation of the gain switched pulse. Since the pump pulse (duration 20 ns) is much shorter than all time constants of the processes of interest the initial conditions simply represent a distribution of excited atomic iodine. This distribution, however, differs from a thermalized Doppler profile. It exhibits a larger mean velocity and a

velocity anisotropy. In [7] it was shown that to a good approximation the velocity distribution can be written as a product of a $|v|$ and a θ dependent term

$$F(\vec{v}, t) = f(|v|, t)\beta(\theta, t) \tag{1}$$

where $v = \sqrt{|\vec{v}|^2}$ is the mean velocity of the gas and θ is the angle between the exciting electric field vector and the propagation direction of the probe laser. The latter in our case is the resonator axis (z-axis) of the iodine laser, i.e., $\theta = 90^o$. Two different relaxation processes can be distinguished, both being collisional. Collisions which change the direction of the velocity vector are responsible for the relaxation of the anisotropy with a characteristic time constant T_b. Collisions changing the magnitude of the velocity (time constant T_v) cause the relaxation to a thermalized Doppler profile with a mean velocity determined by the room temperature. For heavy collision partners it was found that $T_b < (<)T_v$. An exact modeling of the thermalization would require an extensive analysis of the collision processes including the particular properties of the participating particles. As done successfully for the interpretation of the pump-probe experiments in [7] we shall instead describe the thermalization in terms of the two time constants T_v and T_b. To model the evolution of the laser pulse we need to include the interaction with the laser field and its effect on the relaxation.

Let us assume that we must consider $2N+1$ groups of excited iodine of number density N_n ($n = 0, 1, \ldots, N$) with velocity components along the z-axis such that the separation of their transition frequencies is $\Delta\omega_h$. A Doppler velocity profile is characterized by the following distribution of excited iodine with transition frequency $\omega_n = \omega_0 \pm n\Delta\omega_h$

$$N_n^e = N_0 exp \left[-4 \ln 2 \left(\frac{\omega_n - \omega_0}{\Delta\omega_D(t)} \right)^2 n^2 \right] \tag{2}$$

where $\Delta\omega_D \sim \frac{v}{c}\omega_0$. Since the velocity profile is always symmetric, $n = -N, (-N+1), \ldots, N$. For the time dependence of the Doppler width we can thus use the known time dependence of the mean velocity and have

$$\Delta\omega_D(t) = \Delta\omega_{Df} + (\Delta\omega_{in} - \Delta\omega_D) e^{-t/T_v} \tag{3}$$

where $\Delta\omega_{in}$, $\Delta\omega_{Df}$ are the Doppler width of the profile of the "hot" gas immediately after excitation and after complete thermalization with the surrounding, respectively.

At any given instant the number densities obey the following set of differential equations describing the relaxation towards a Doppler profile

$$\frac{dN_n}{dt} = \frac{N_n^e - N_n}{T_b} \tag{4}$$

where N_n^e is given by Eq. (2). Under laser action we need to distinguish between excited and ground state iodine, each group undergoing similar relaxation processes. While excited iodine is generated

at many different transition frequencies ground state iodine is only produced by simulated emission at the transition frequency ω_0. Collisional relaxation processes then cause a redistribution towards a thermalized Doppler profile.

3.2 Laser model

To model the operation of a photolytic iodine laser the typical rate approximation approach was chosen. In accordance with the experiments we assume single longitudinal mode operation. Therefore at a certain time t only atoms with resonance frequency ω_0 interact directly with the laser field (photon flux F). The set of rate equations can then be written as

$$\frac{d}{dt}N_0 = \sigma_{0\ell}F(N_{0\ell} - N_0) + \frac{1}{T_b}(N_0^e - N_0) \tag{5}$$

$$\frac{d}{dt}N_{0\ell} = \sigma_{0\ell}F(N_0 - N_{0\ell}) + \frac{1}{T_b}(N_{0\ell}^e - N_{0\ell}) \tag{6}$$

$$\frac{d}{dt}F = \sigma_{0\ell}cF(N_0 - N_{0\ell}) - \frac{F}{T_R} \tag{7}$$

$$\frac{d}{dt}N_n = \frac{1}{T_b}(N_n^e - N_n) \qquad n = 1, 2, \ldots, N \tag{8}$$

$$\frac{d}{dt}N_\ell = \sigma_{0\ell}(N_0 - N_{0\ell}) \tag{9}$$

where $\sigma_{0\ell}$ is the amplification cross section, $N_{0\ell}$ is the number density of de-excited atoms of transition frequency ω_0, T_R is the resonator decay time, N_n^e is the number density of atoms if distributed according to a Gaussian profile of width $\Delta\omega_D$, and c is the velocity of light. The quantity N_ℓ describes the total number of iodine atoms in the ground state. The total number of excited iodine atoms is

$$N_{total}(t) = N_0(t) + 2\sum_{n=1}^{N} N_n(t) \tag{10}$$

and the equilibrium numbers can be calculated from

$$N_0 = N_{total}\left[1 + 2\sum_{n=1}^{N} e^{-n^2(\Delta\omega_h/\Delta\omega_D)^2}\right]^{-1} \tag{11}$$

and

$$N_n = N_0 \exp\left[-\left(\frac{\Delta\omega_h}{\Delta\omega_D}\right)^2\right] \tag{12}$$

for $n = 1, 2, \ldots, N$. After each integration step these equilibrium numbers have to be determined with the modified N_{total} and $\Delta\omega_D$ according to Eqs. (10) and (3), respectively. The same procedure must be applied for the atoms in the ground state. Their total number is given by the total number
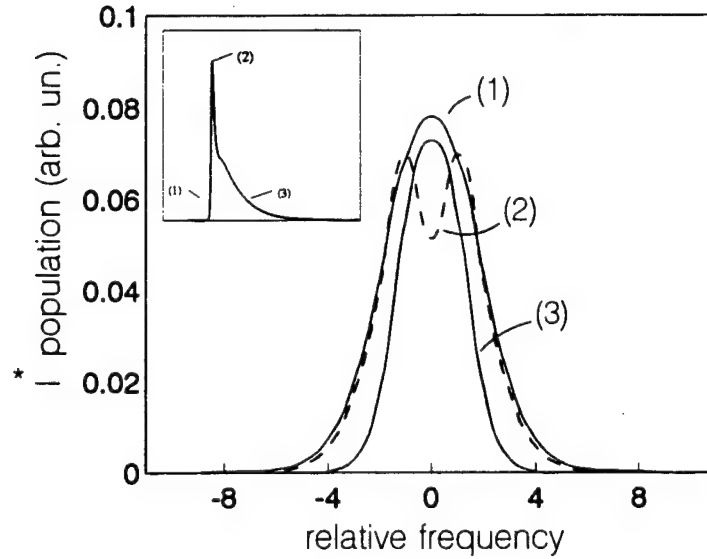
Figure 7: Calculated gain profile vs frequency as seen by different parts of the iodine laser pulse. The inset shows a typical iodine pulse and the corresponding times.

of de-excitation events, where each event results in the occupation of $N_{0\ell}$ only. Note that we can neglect spontaneous emission for this process because of the smallness of the A coefficient ($\approx 10s^{-1}$) of the magnetic dipole transition of iodine. A small number of (spontaneously emitted) photons are injected at $t = 0$, though, to induce the build up of the laser pulse. The integration of the preceding system of differential equations (5)-(9) was performed using a Runga-Kutta routine.

3.3 Results of the computer simulation and comparison with the experiment

Figure 7 shows the occupation numbers in the excited state as a function of the transition frequency for different times after excitation. The inset shows the corresponding iodine laser pulse. Curve (1) describes the distribution at the pulse leading edge (at 10% of the peak value). The non Gaussian distribution here has already relaxed into a Gaussian (relaxation time T_b) which still corresponds to a velocity much larger than the equilibrium value. At the pulse peak (curve (2)) spectral hole burning is clearly visible which causes a relaxation from excited groups of atoms into those in resonance with the laser field. This means a repumping of the laser transition. The relaxation of the magnitude of the velocity (relaxation time T_v) is another process which results in an effective repumping. As the Doppler profile narrows the number of excited atoms with transition frequency ω_0 increases to a constant that depends on the total number of excited iodine atoms and room temperature as described by Eqs. (2) and (3). (if we neglect de-population by the laser

pulse). Curve (3) shows the gain profile seen by the trailing edge of the pulse (at 10% of the peak value). A comparison of curves (1) and (2) illustrates our discussion. Due to the laser action the area under curve (3) is smaller. Of course, for the actual laser action, the difference $(N_0 - N_{0\ell})$ is of importance. Due to the combined action of stimulated emission and the thermalization of ground state iodine, $N_{n\ell}$ is constantly increasing with time; therefore, the qualitative features of the gain can be explained in terms N_n.

As seen in the experimental section, at low CF_3I pressure the laser pulse exhibits a characteristic shape-a sharp rise time and a distinct shoulder. The pulse duration and decay time are significantly longer than the resonator life time. Figure 8 compares preliminary results of measured and calculated pulse profiles. From the experiment the resonator life time and the number density of excited iodine can be estimated. For the homogeneous line width we assumed $\Delta\omega_D = 60$ MHz. The only remaining input parameter for the simulation are the two relaxation times T_b and T_v. For C_3F_7I and a gas pressure of 0.2 Torr values of 1 μs and 2.5 μs have been measured [7]. For the simulation we assumed that both relaxation constants are inversely proportional to the pressure and that their ratio is constant. all other input parameters were not changed. From the simulation it becomes obvious that the measured features are a consequence of the collisional relaxations. The shoulder at low pressure is a result of the hole burning and a fast refill of excited iodine at the laser frequency owing to the short relaxation involving T_b. With increasing pressure and thus shorter T_b the second maximum vanishes and the main peak simply broadens. The decay of the iodine pulse is essentially controlled by the $|v|$ relaxation proceeding more slowly.

Another important feature is the delay between excimer pump pulse and the occurrence of the iodine pulse. The decrease of the delay with increasing buffer gas pressure is also a result of the faster relaxation. Figures 9 (a) and (b) show the delay and the decay time as function of the pressure (here as function of T_b while keeping T_b/T_v constant). The behavior of the curves describes the experimental findings qualitatively.

For a more complex (future) quantitative comparison of simulation data and experiments we plan to implement the following in the discussion: The addition of a buffer gas (BF) requires the introduction of two different relaxation rates. If the buffer gas pressure is changed their value will change accordingly while their ratio remains. In the simulation we then need to use

$$\frac{1}{T_{b,v}} = \frac{1}{T_{b,v}(CF_3I)} + \frac{1}{T_{b,v}(BF)} \tag{13}$$

where only the second summand changes with a change in buffer gas pressure. The different effect of adding CF_3I or buffer gas on the measured pulse characteristics suggests the necessity of this approach.

Since the photolytic iodine laser is a relatively simple system with most parameters well documented we believe that the performed experiments can be used to determine relaxation constants
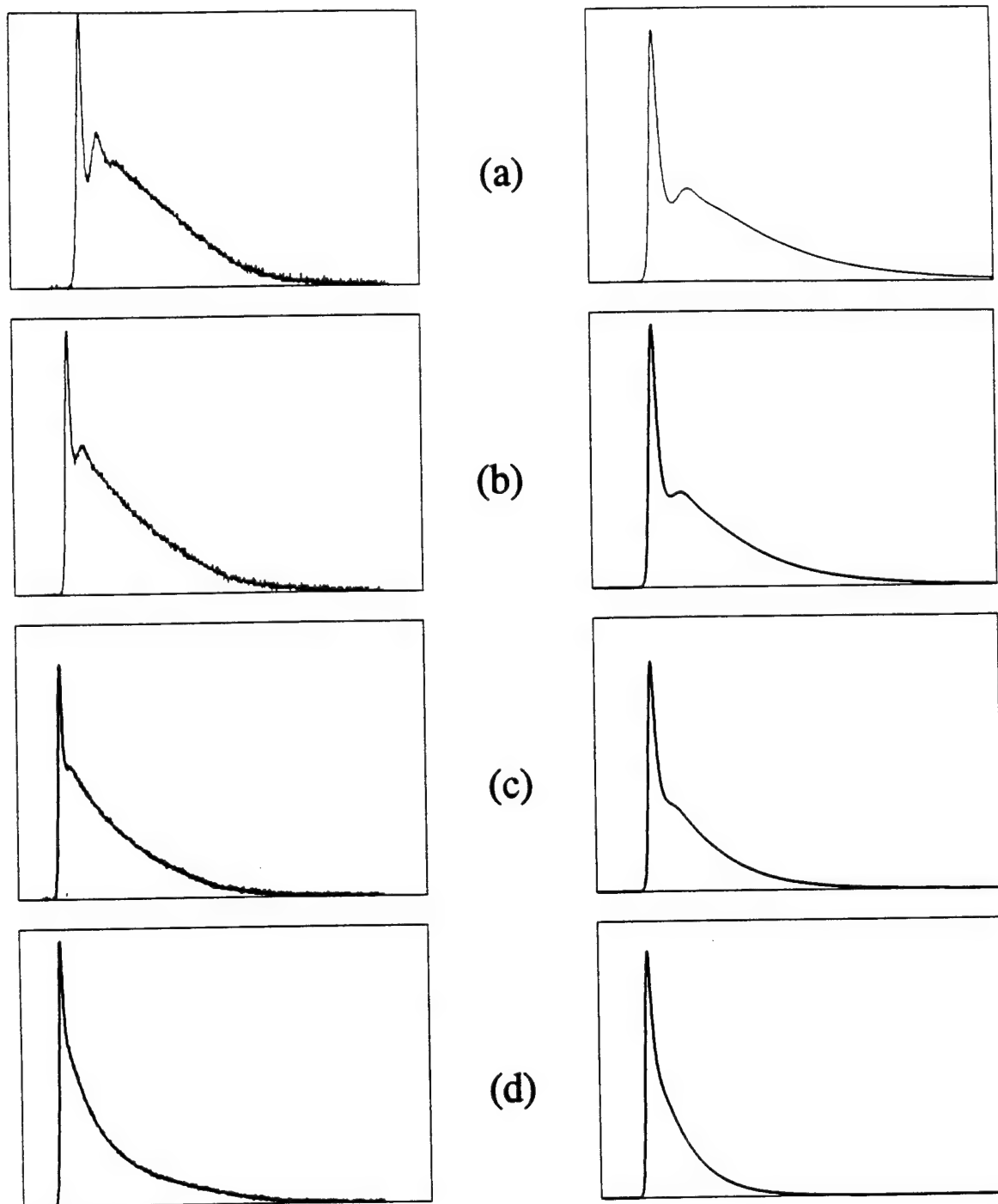
Figure 8: Measured (left) and calculated (right) pulse profiles. The experimental data corresponded to 0.6 Torr CF$_3$I plus (a) 0 Torr, (b) 0.3 Torr, (c) 0.6 Torr, and (d) 0.9 Torr Xenon buffer gas. For the numerical simulation we assumed $T_b/T_v = 0.38$ and changed T_b from 650 ns to 260 ns from top to bottom in increments equal to the total gas pressure change. The length of the time axis is 6 μs for all the profiles.
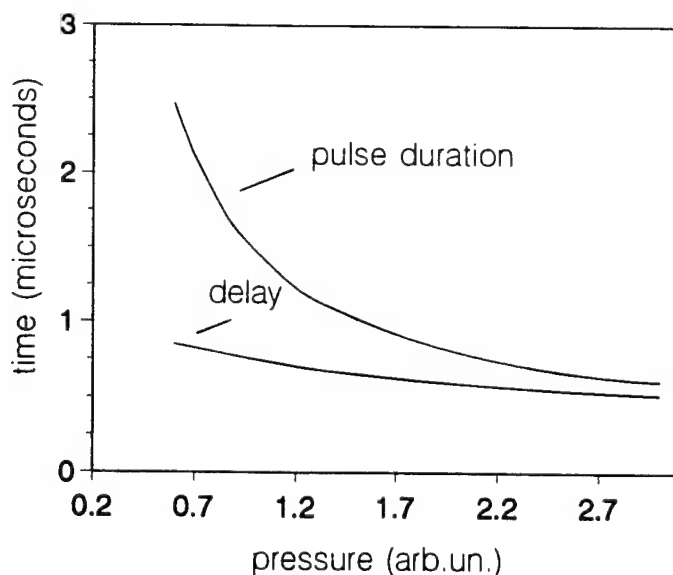
Figure 9: Delay between pump pulse and the ten percent of peak value of the iodine pulse and decay time of the iodine pulse as a function of the buffer gas pressure. In the simulation we actually changed the values of T_v, T_b as described in the text.

as function of buffer gas composition and pressure. This new type of "spectroscopy" would not require an additional probe laser.

4 Summary

The dynamics of gain switched iodine lasers at low pressure is determined by collisional relaxation processes involving a change in the direction of the velocity vector and a change of its magnitude. Both relaxation constants depend on the gas pressure and the type of buffer gas (if any). A comparison of experimental data with a computer simulation allows one to determine these relaxation rates. Important pulse features that need to be analyzed and modeled are the pulse build-up time, the pulse shape and its decay time. The two relaxation mechanisms provide a repumping mechanism of a single mode iodine laser which manifests itself in the occurrence of a distinct shoulder and a second pulse peak at low gas pressure.

References

[1] M.A.Kelly, J.K.McIver, R.F.Shea, and G.D.Hager. *IEEE J. Quantum Electron.*, QE-27:263, 1991.

[2] G.Brederlow, E.Fill, and K.J.Witte. *The high-power iodine laser.* Springer, 1983.

[3] L.A.Schlie and R.D.Rathge. *Opt.Commun.*, 66:289, 1988.

[4] C. M. Ferrar. *Appl. Phys. Lett.*, 12:381, 1968.

[5] H.J.Baker and T.A.King. *J.Phys.D:Appl.Phys.*, 8:L31, 1975.

[6] A.S.Grenishin, V.M.Kiselev, and T.N.Koltlikova. *Opt.Spectrosc. (USSR)*, 52:212, 1982.

[7] J. I. Cline, C. A. Taatjes, and S. R. Leone. *J. Chem. Phys.*, 93:6543, 1990.

[8] R. Schmiedl, H. Dugon, W. Meier, and K. H. Welge. *Z. Phys. A*, 304:137, 1982.

CHARACTERIZATION OF POLAR PATCHES: COMPARISON OF MULTIPLE SENSORS

Gary S. Sales
Associate Professor
Department of Electrical Engineering
University of Massachusetts Lowell
1 University Avenue
Lowell, Massachusetts 01854

CHARACTERIZATION OF POLAR PATCHES: COMPARISON OF MULTIPLE SENSORS

Gary S. Sales
Associate Professor
Department of Electrical Engineering
University of Massachusetts Lowell

Abstract

Polar patches are relatively intense ionospheric structures formed on the day side of the high latitude polar region and transported across the polar cap in the anti-sunward direction towards the nighttime auroral oval. The basics of these patches has been known for some time, initially characterized using vertical ionosondes. More extensive investigations into the character of the three dimensional structure of these patches required more sophisticated systems such as the incoherent scatter radar that can scan through the patches and optical systems that photograph the entire sky from horizon to horizon at a wavelegth that distinguished the patches from the less intense background emissions. To be able to continue these investigations at locations where no radar is available is the goal of this project. Modern digital HF sounders, besides measuring the vertical electron density profile, can also detect and determine the location of small scale ionospheric irregularities which in turn map the overall structure of the large scale patches. This study demonstrates the capabilities of these HF sounders.

CHARACTERIZATION OF POLAR PATCHES: COMPARISON OF MULTIPLE SENSORS

Gary S. Sales

Introduction

Polar patches are "blobs" of daytime F-region plasma transported in the anti-sunward direction over the polar cap from the cusp region towards the nighttime aurora. These patches are rather large structures, of the order of several hundreds of kilometers, drifting at speeds of 500 to 1000 m/s in the polar ionospheric circulation. They originate as a result of instabilities in the ExB flow pattern near the cusp. Internally these patches appear to be highly structured when observed by transionospheric signals from satellites. These signals fade rapidly and deeply when the satellite signal traverses a patch region. These satellite scintillation data indicate a small scale field aligned irregularity size of the order of several kilometers to tens of kilometers within the larger patch structure. These patches have mean densities considerably above the ambient nighttime peak F-region density within the polar cap. Typically the maximum plasma frequency within the patch is in the range of 5 to 7 MHz compared to the ambient nighttime background density of 2.5 MHz at the peak altitude near 400 km.

This study involves data collected at Sondrestrom Fiord (67°N, 309°E) in Greenland during the Rodeo II campaign in October, 1989. The rather steady flow of patches over the site was observed by optical means, specifically, the all-sky photometer operating at 6300 A and by the incoherent scatter radar (ISR) measuring the ionospheric electron density and the plasma drift velocity over the site. A vertical incidence digital ionospheric sounder operated in both the ionogram and drift modes during this campaign. The data collection schedule during a 15 minute period of this campaign is shown in Figure 1.

The first goal of this project is to compare the description of these polar ionospheric patches as viewed by these different sensors and then to combine them to generate a more complete characterization. Secondly, it is expected that a detailed characterization of complex ionospheric structure can be obtained without the use of the incoherent scatter radar. There are few of these radars and to the extent that modern HF sounders can be used to

characterize polar patches or other features such as equatorial plumes, a new and effective diagnostic tool could become widely available.

Data Collection

Incoherent Scatter Radar

The Sondrestrom incoherent scatter radar operated in a special scanning mode during the Rodeo II campaign. The radar measures the local electron density and plasma drift velocity, both of which are important to this study. Three parallel scans, taking a total of 10 minutes, were executed in a pattern shown in Figure 2. The direction of this scan was determined by the orientation of the auroral arc siting to the south of the station at that time. The requirement was that the scan be approximately perpendicular to the length of the auroral arc. During these auroral measurements, above the station, plasma patches were easily observable in the F-region of the ionosphere. The three electron density cuts, from 0006 UT to 0017 UT through these patches are shown in Figure 3. The long integration times ($\approx$ 10s) and long radar pulses results in relatively smooth electron density profiles.

To construct a 3-D model of the electron density in the region over Sondrestrom, the three radar cuts at the peak altitude of 375 km where extrapolated in latitude, longitude and in altitude. The structure sizes seen for these patches are typical and the entire pattern is drifting with a measured speed of 600 m/s along the path of the scan in the anti-sunward direction, approximately from north to south. Because it takes ten minutes from the beginning of the scan to the end, a distortion, primarily in the orientation of the long axes of the drifting patches relative to the radar scan, is introduced.

To correct this distortion it was decided to restore the patches as they would appear at the time corresponding to the center of the scan, in this case 0011 UT. Using the measured drift speed of 600 m/s, the measured electron densities were moved forward or backward depending on whether the measurement was before or after 0011 UT. The corrected electron density profiles are shown in Figure 4.

All-Sky Photometer (6300 A)

Operating at Sondrestrom is a unique scanning photometer. Of primary interest for this study is the all sky photographs of emissions from the F-region, at 6300 A, primarily from the intense patches and from the arurora to the south. During Rodeo II the actual goal of the campaign was to investigate the structure of these auroral arcs. The patch data collection was somewhat fortuitous even though the optical system was attenuated to prevent the very strong arc emission from overloading the system. This made the weaker patch emissions more difficult to observe.

The all sky photographs at 6300 A were made at the rate of one per minute and each was exposed for 1s. With digital processing these optical data can be superimposed on a geographical map as shown in Figure 5. These data, for times around the ISR scan, have been placed into a latitude/longitude grid, the same as the radar data and all the other data used in this study. These optical data appear to be very similar to the radar data and serve to confirm the corrected patch structure derived from that data.

Digisonde Vertical Ionograms

The vertical soundings made at Sondrestrom during Rodeo II were made every 5 minutes using the University of Massachusetts Lowell Digisonde 256 system. In the vertical ionogram, the Digisonde is able to display the direction of the arriving signals in terms of twelve directions and vertical. The relevant ionograms from 0009 UT and 0014 UT are shown in Figure 6. The dark traces represent "overhead" signals while the grey areas are off-vertical returns out to about 50° from zenith.

For these data several off-vertical returns are observed at both times. As is shown in the analysis below, these signals represent the oblique reflections from several of the patches. The "vertical" trace moves significantly between the two ionograms separated in time by only 5 minutes.

Digisonde Sky Maps

During the approximately four minutes between the end of one and the beginning of the next vertical ionogram the Digisonde system goes into a fixed frequency drift mode, measuring the arrival angles and Doppler

frequency shifts associated with separate reflecting sources in a structured ionosphere. Satellite scintillation measurements in the polar cap show that these patches are highly structured with field aligned irregularities in electron density. Analysis of the scintillation measurements indicate an irregularity structure size of the order of 10 km across the field lines. This scale size is sufficiently large to strongly reflect HF radio waves in the frequency range of 3 to 6 MHz when the radio wave is orthogonal to the irregularity.

The Digisonde system processes the signals arriving on the array of receiving antennas to determine the amplitude and phase of each spectrum line and then to locate in angle the source associated with that line. Typically, the Digisonde automatically selects a frequency for the drift mode from the previous vertical ionogram usually near the lower frequencies of the F-trace. Most of the data presented here is in the frequency range of 3 to 5 MHz and at the shortest time delay associated with the F-layer trace.

The identified sources from each coherent integration time (10 s.) and from two adjacent range gates covering about 20 km are combined to form a sky map. Figure 7 shows composite skymaps made at 0008, 0011 and 0014 UT, generated by combining 3 sequential sky maps at the same sounding frequency and range gate. At 0008 UT there appears to be two clusters of sources, a weaker one NW of the site and a second one towards the SE. At the other two times only one cluster of sources appears to the NW of the sounder site.

HF Ray Tracing Simulation

A new ray tracing code, called HASEL, developed by Coleman (1993) with the ability to input an ionospheric model based on a latitude/longitude/altitude grid with any desired spacings, is used here to simulate expected vertical ionograms and sky maps showing the location of sources connected with the patches as seen by the ISR and the all-sky optical system. Once this capability is demonstrated, then it should be possible to invert this process and use the sounder data to characterize these structures.

As discussed in Section 2.1 the 3-D corrected electron density profiles were used to develop a simple analytical model with Gaussian ellipsoidal patched that closely match the densities measured by the ISR. These corrected

profiles are referenced to the mid-time of the radar scan, 0011 UT. HASEL can be operated with or without the magnetic field effect on the refractive index. For this study the magnetic field was not used to calculate ordinary and extraordinary modes, but is used to determine when each ray path becomes perpendicular to the earth's magnetic field line. This test for orthogonality with the earth's magnetic field is a simple method to simulate the small scale irregularities that lie within the patches. The basic assumption used here is that these small scale irregularities pervade the patch and when the ray path is perpendicular to the field line it encounters a filed aligned irregularity and a reflection occurs, returning the wave along the same ray path back to the sounder. The inclusion of 10 km scale irregularities into the model is possible but would make the ray tracing almost impossible, requiring a grid spacing small compared to 10 km. The grid spacing used here is 0.2° in latitude and longitude (≈ 20 km) and 5 km in altitude.

Finally, a test is set to insure that only orthogonalities within a patch are used to simulate the observed structure. Here the requirement is that at the orthogonality point the electron density is more than 20 % above the background ionosphere (i.e. the ionosphere without any patches) at that altitude. The background ionosphere is determined using the ISR data in areas where no patches are observed.

Using the 0011 UT patch model (Figure 4), ray tracing at 3 MHz, the frequency selected by the automatic system in the Digisonde drift mode, was carried out. This ray tracing was done for azimuths from 0° to 330° in steps of 30° and at each azimuth for elevation angles from 45° to 89° in steps of 2°. When additional information was needed, some ray paths at intermediate azimuths were calculated.

For comparison with the actual recorded sky amps only the shortest delayed points can be used since the Digisonde 256 selects only the bottom of the F-layer trace for data recording. To compare the simulations for the other sky map times additional ionospheric models would have been necessary. To avoid the extrapolation of the 0011 UT model to other times a simple method was devised where the location of the sounder location was moved in the direction opposite to the patch drift velocity, corresponding to 3 minutes

earlier that 0011 UT and then 3 minutes later. Then the ray tracings where repeated corresponding to the skymap times shown in Figure 7.

To simulate the vertical ionograms, the same ray tracing calculations, for azimuth and elevation, were carried out for the three times at all frequencies between 3 MHz and 8 MHz in 0.5 MHz steps. These results are shown in Figure 8 in a standard ionogram format, i.e., time delay (km) vs. frequency. Here, for clarity, the twelve azimuths are separated in the direction of increasing frequency to make it possible to distinguish orthogonality points corresponding to the distinct patches.

Analysis

Having used the incoherent scatter radar electron density data to establish the model and the optical data to confirm the correction to the profiles, the only remaining comparisons are with the vertical ionogram and the sky map data. For the sky map, the ray tracing simulation has been limited to the minimum time delay sources and contoured to outline the source region. Figure 9 shows these contours superimposed over the three sky maps reproduced from Figure 7. Only at 0008 UT does the simulation indicate that two patches can be seen at the same time. Returning to Figure 4, the sounder is located between the two large patches and both lie at about the same time delay from the sounder. The cluster of points from the actual sky map fall well within the contour indicating the correctness of the simulation.

Considering the apparent motion of the station, the sounder moves under the edge of the northern patch at 0011 UT and the situation changes. Now at the sky map frequency and for minimum time delay only the northern patch is visible both in the actual sky map and in the simulation and again the match is good. Finally, at 0014 UT, the station moves well under the northern patch as shown in Figure 4 and again only the northern patch is visible in the data and in the simulation. In each of these cases, the simulations were done at the frequency used to generate the actual sky maps. As discussed later, there are better frequencies that do a more complete job of mapping the patches in the sounder area. In all of these comparisons the simulations are good in terms of the location of the reflection sources and

and the time delay for each situation, within the accuracy of the system, i.e., 15 km or about one pulse width.

For the vertical ionogram comparison, the simulated ionograms closest to the actual at 0009 and 0014 UT were examined to determine which of the points correspond first to the vertical trace. This was done by finding those rays near vertical rays that come closest to the sounder location. It was assumed that these sources make up the vertical trace. For the three simulated ionograms, at 0008, 0011 and 0014 UT, these vertical traces are marked. These are to be compared with the vertical traces indicated in black on the actual vertical ionograms shown in Figure 6.

Finally, the ray tracing simulations were examined to identify which orthogonal sources belong to each of the major patches. Although in the skymap usually only one patch is visible because of the way the sounder measures only at the shortest time delay in the drift mode while in the ionogram mode all delays are measured and shown and several patches can be observed. When these results are compared to the actual ionograms, the structure of the ionogram becomes clearer. The spur which extend beyond the vertical trace clearly result from the two large patches in the vicinity of the sounder. This comparison reinforces the suggestion that higher frequencies can do a better job of mapping the local patch structure in the drift mode than the 3 or 4 MHz frequencies that are selected by the current system. Figure 10 shows the results of ray tracing at 6 MHz at 0011 UT. Here the elctron density contour from Figure 4 have been overlayed for comparison. At this frequency the mapping of the patch structure is quite good.

Summary

Starting with the measured electron density profiles it is possible, using 3-D ray tracing, to synthesize the sounder measurements including the detailed structure of the vertical ionograms and the source location of the Digisonde sky map system. The source simulation at the higher frequencies clearly shows the overall structure of these polar patches and it is possible to determine the internal character of small scale irregularities as well as the overall electron density distribution. The newer University of Massachusetts Lowell DPS system has improved capabilities and will make it possible to map polar patches as well as other ionospheric structure without the necessity of an ISR at the same location.
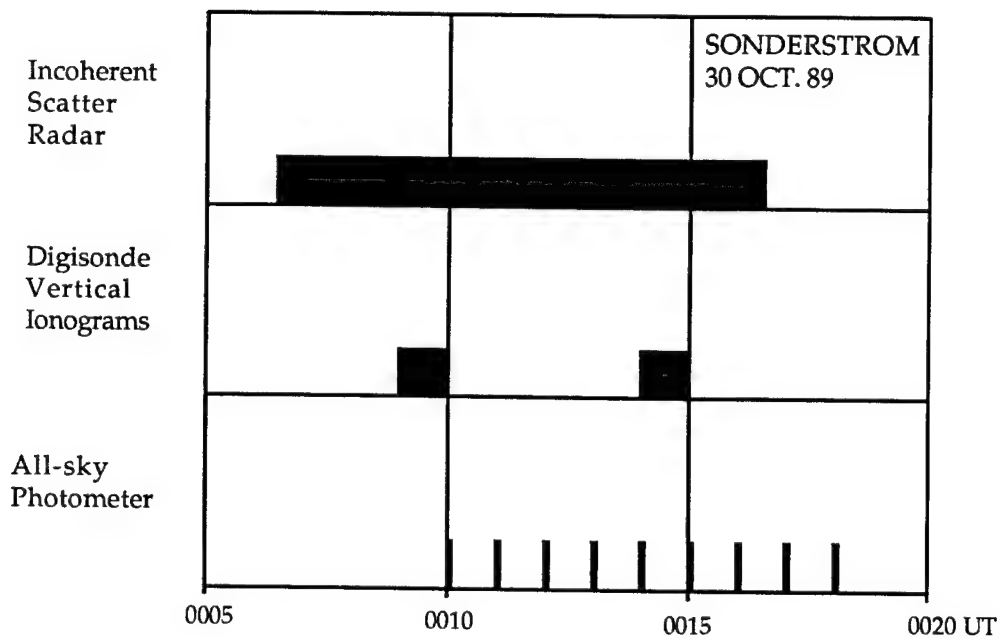
Figure 1. Survey of data collected from 0005 UT to 0020 UT on 30 Oct. 1989
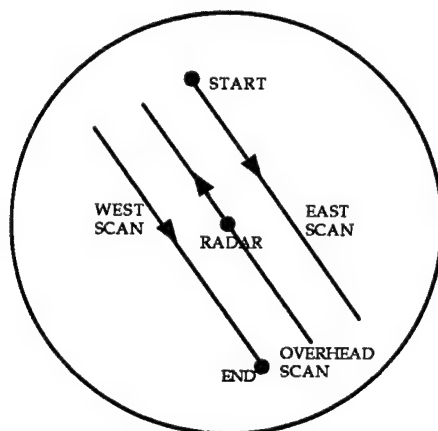
Figure 2. Scan pattern of incoherent scatter radar at Sondrestrom, Greenland during the Rodeo II campaign; 30 Oct. 1989

30-10
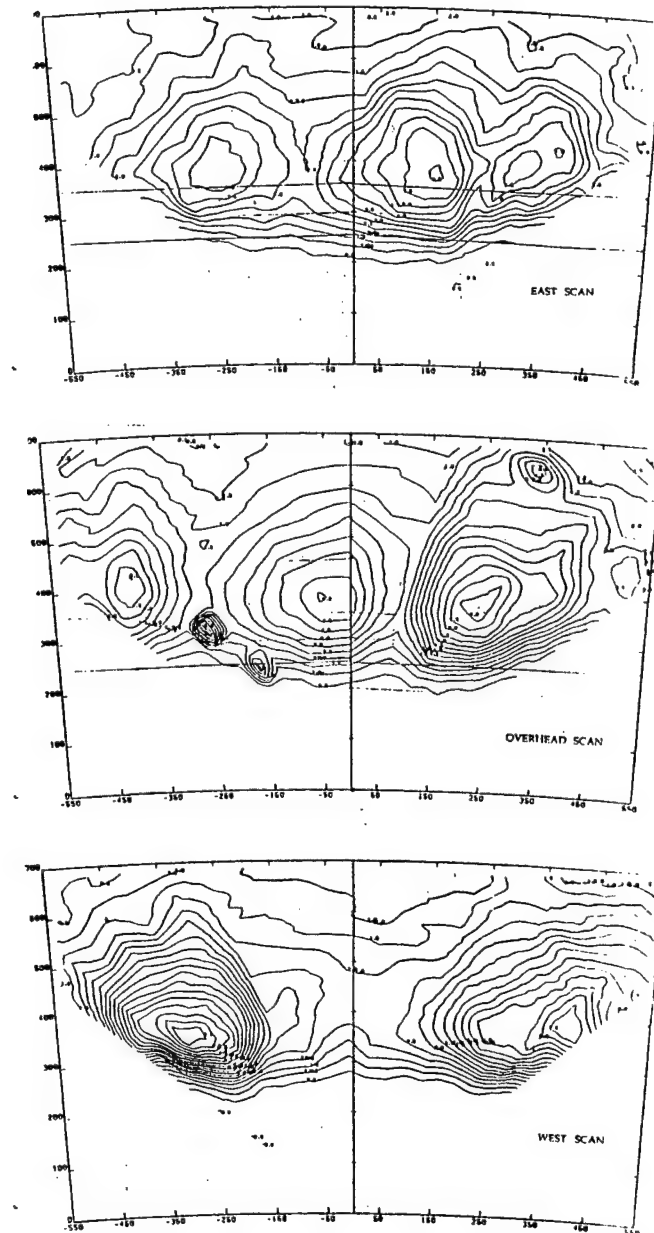
Figure 3. Electron density vertical profile scans made by the ISR (0006-0017 UT)
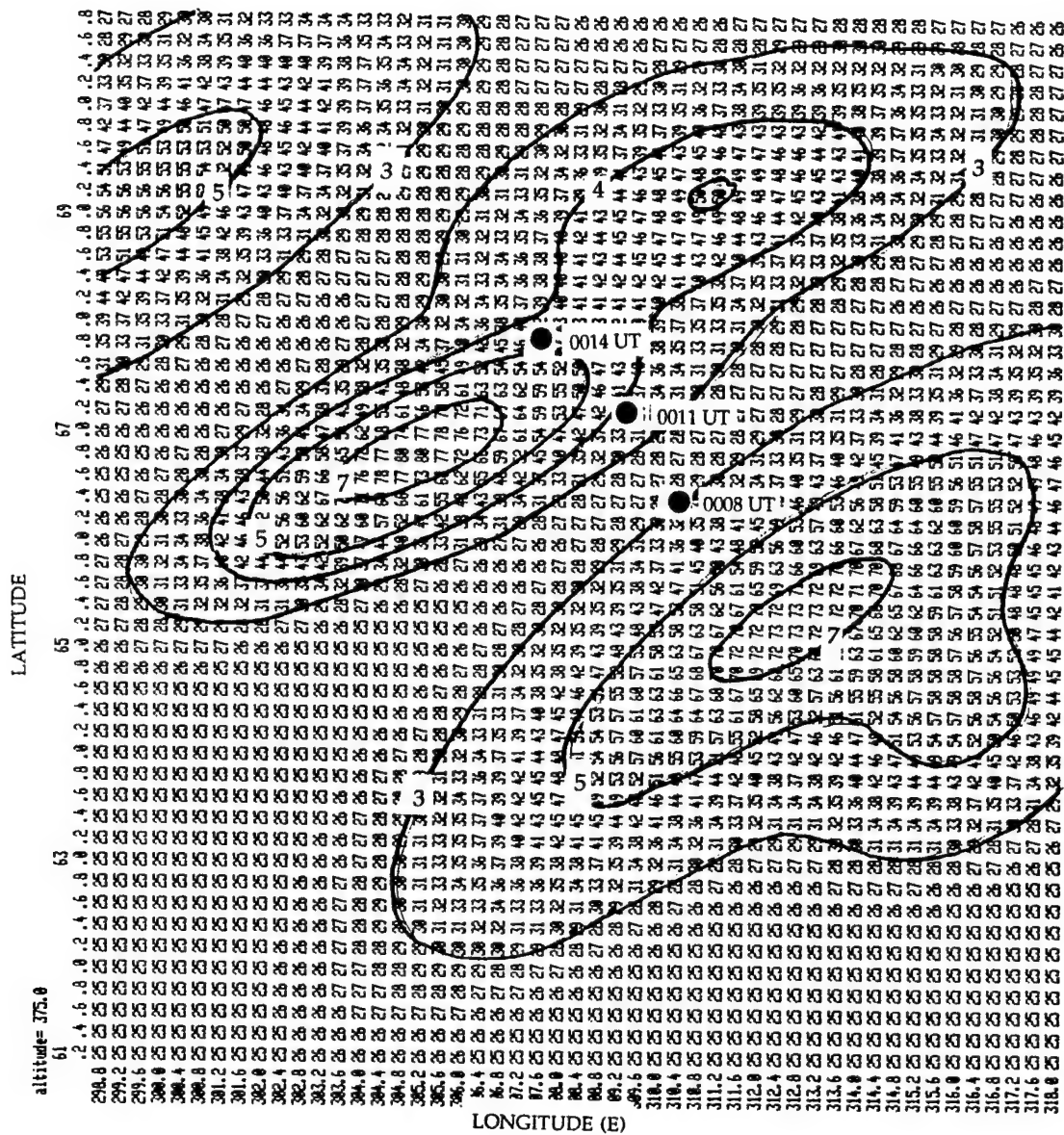
Figure 4. Electron density contours (x 10-5) at 375 km corrected for patch drift
(0011 UT)
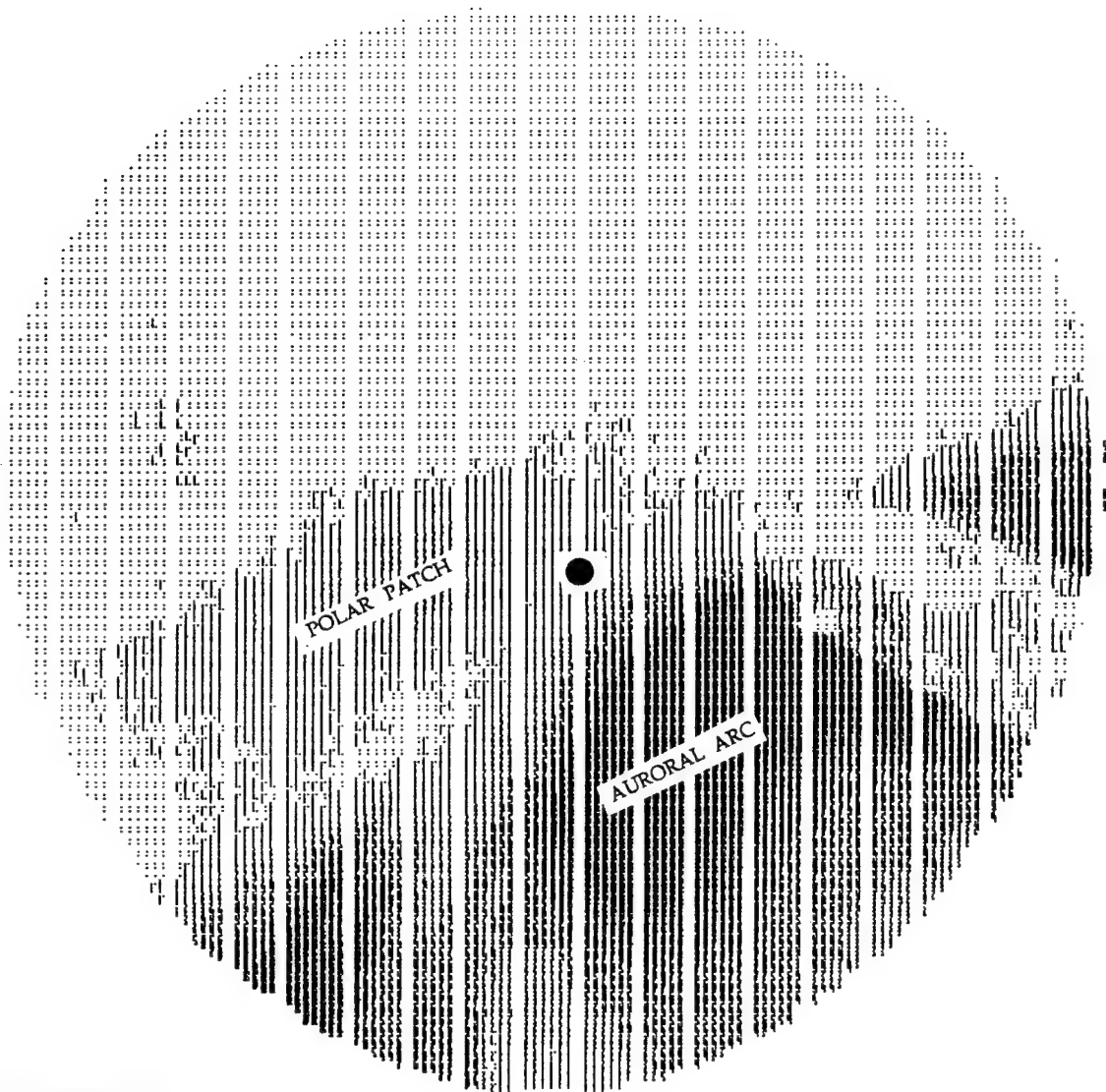
Figure 5. All sky photometer image at 6300A ingeographical coordinates
(0011 UT)
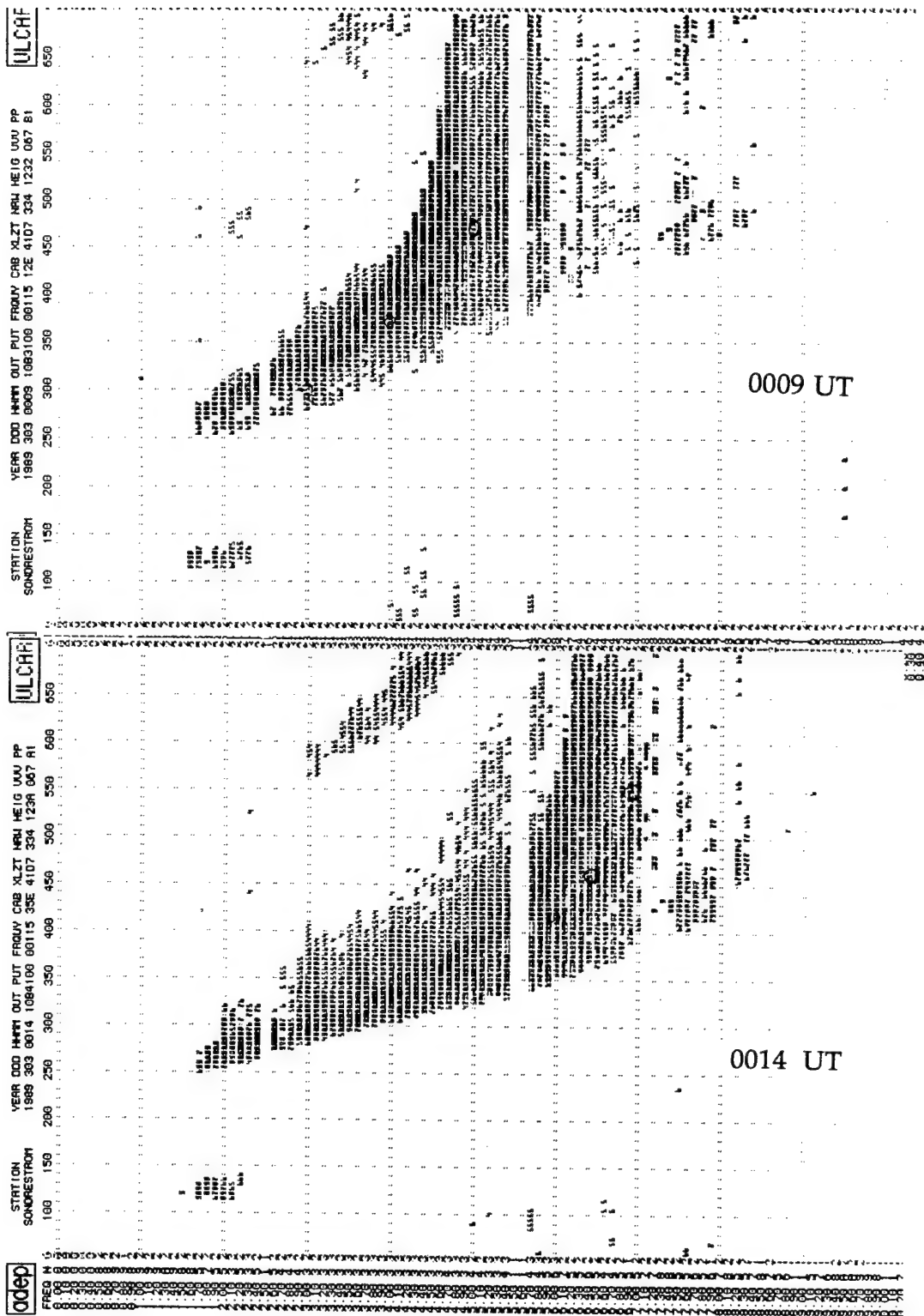
Figure 6. Vertical Digisonde ionograms at 0009 and 0014 UT

Figure 7. Composite Digisonde skymaps made at 0008, 0011 and 0014 UT.
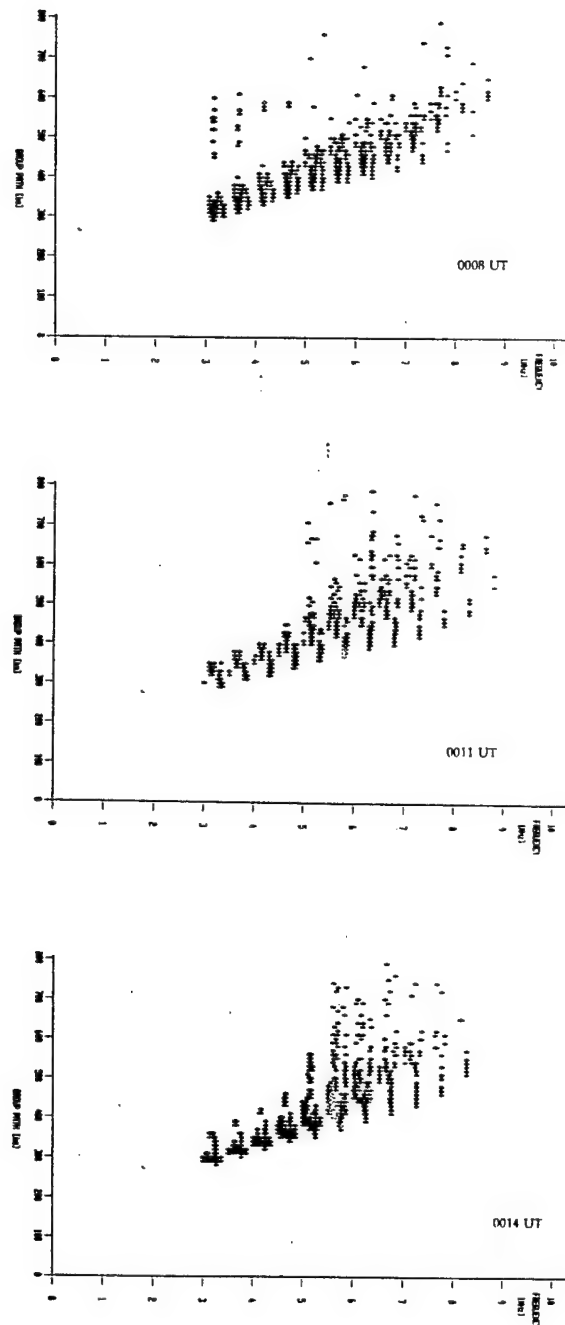
Figure 8. Synthesized vertical ionograms at 0008, 0011 and 0014 UT
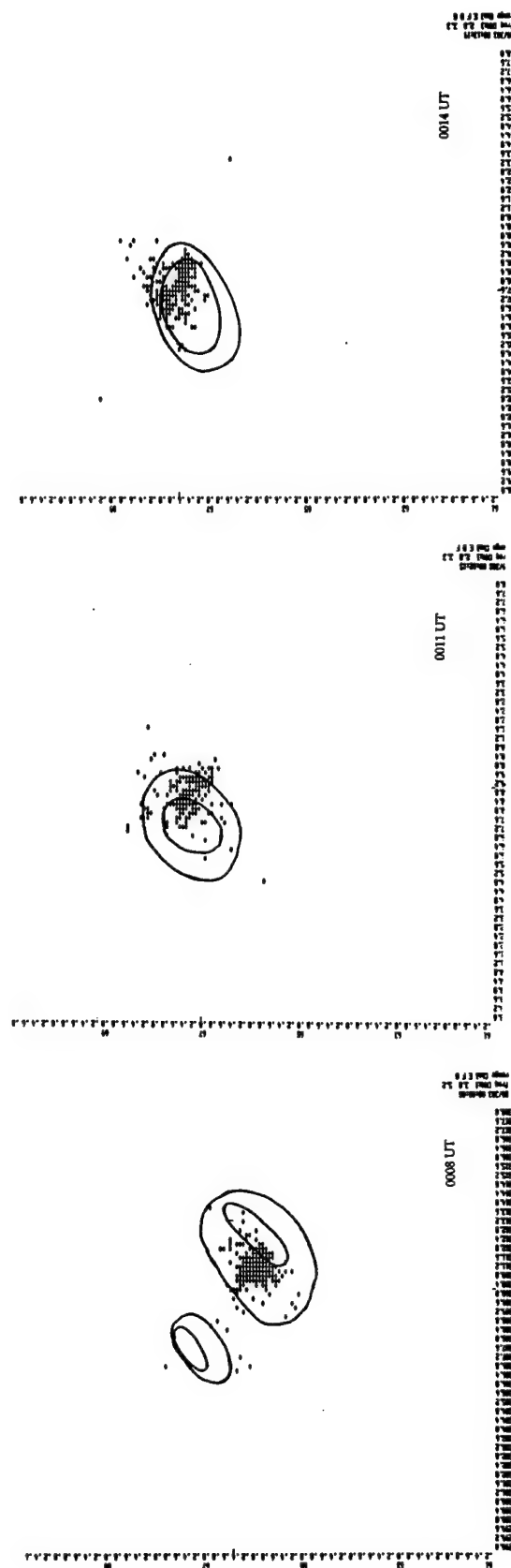
Figure 9. Sky maps at 0008, 0011 and 0014 UT overlayed with the simulated maps using ray tracing.
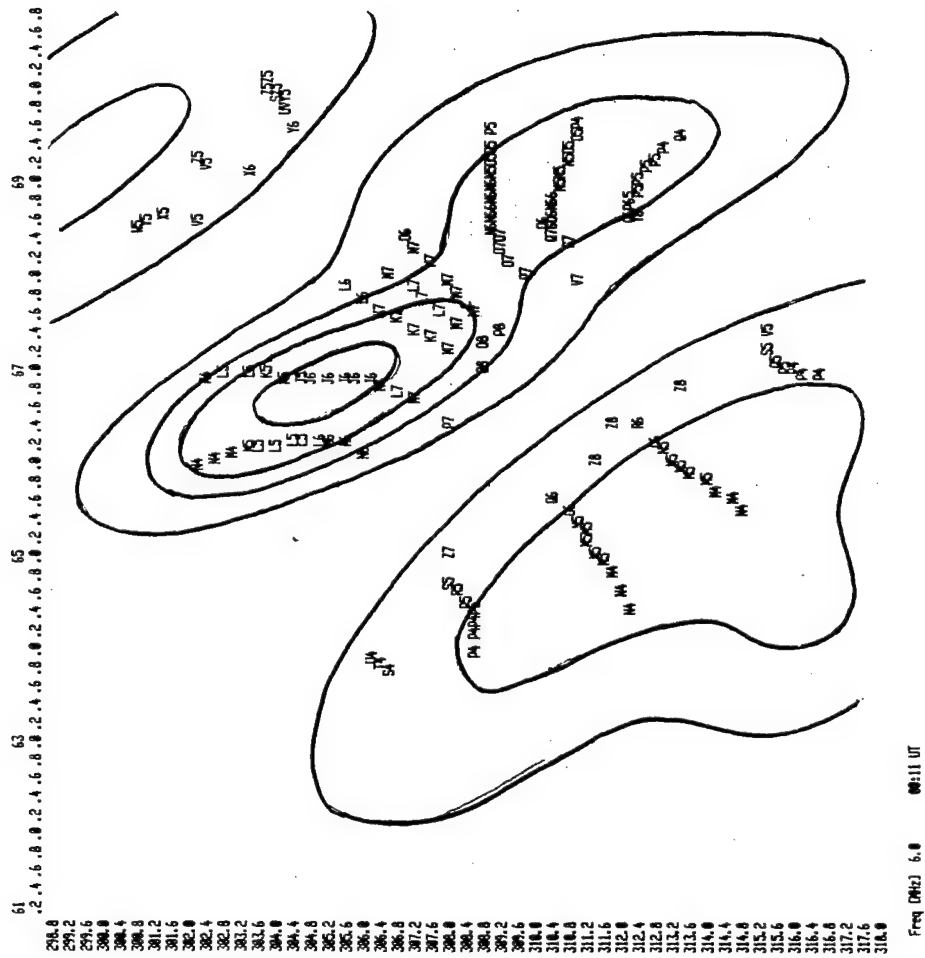
Figure 10. Simulated 6MHz skymap showing mapping of the patch structure.

A STUDY OF ACTIVE CONSTRAINED LAYER DAMPING TREATMENTS ON COMPOSITE BEAMS

I. Y. (Steve) Shen
Assistant Professor
Department of Mechanical Engineering

University of Washington
Seattle, WA 98195

A STUDY OF ACTIVE CONSTRAINED LAYER DAMPING TREATMENTS ON COMPOSITE BEAMS

I. Y. (Steve) Shen

Assistant Professor

Department of Mechanical Engineering

University of Washington

Seattle, WA 98195

Abstract

This research has two objectives. The first objective is to develop a mathematical model to predict bending, twisting, and axial vibration response of a composite beam with active constrained layer (ACL) or intelligetn constrained layer (ICL) damping treatments. The second objective is to verify the concept of active constrained layer through laboratory experiments and to evaluate the strength and weakness of this new technique.

To achieve the first objective, an ICL composite beam model is obtained by integrating the existing ICL composite plate model proposed by Shen (1994a). When the plate width (along the x-axis) is much smaller than the plate length (along the y-axis), integration of the ICL composite plate equations and linearization of displacement fields with respect to x will lead to a set of equations that couples bending, tosional, and axial vibrations of a composite beam. The equations of motion and associated boundary conditions are normalized and rearranged in a state-space matrix form, and the vibration response is predicted through the transfer function approach developed by Yang and Tan (1992). A numerical example is illustrated on a composite beam with bending-torsion coupling stiffness. Numerical results show that ICL damping treatments may or may not reduce coupled bending and torsional vibrations of a composite beam simultaneously. When the deflection is fed back to actuate the ICL damping treatment, a sensitivity analysis shows that only those vibration modes with significant bending response will be suppressed simultaneously with their torsional components.

To achieve the second objective, two different ICL setups are tested on a composite beam without bending-torsion coupling. Damping performance of both ICL setups agree qualitatively with existing mathematical models and experimental results obtained from other researchs. The damping performance, however, is not optimized due to the availability of materials and their dimensions in the laboratory. An optimization strategy needs to be developed to facilitate design of ACL damping treatments with maximized damping performance.

A STUDY OF ACTIVE CONSTRAINED LAYER DAMPING TREATMENTS ON COMPOSITE BEAMS

I. Y. (Steve) Shen

Introduction

Active constrained layer (ACL) or intelligent constrained layer (ICL) damping treatments are hybrid damping designs that integrate both active and passive dampings through constrained layer treatments. A typical design of ICL (see Figure 1(a,b)) consists of a viscoelastic shear layer sandwiched between a piezoelectric constraining layer and the vibrating structure. In addition, the vibration of the structure is fed back to regulate the axial motion of the piezoelectric constraining layer. The vibration energy is then dissipated through cyclic shear deformation of the viscoelastic layer. In this design, the active component will provide adjustable and significant damping, whereas the passive component will enhance gain and phase margins, reliability, practicality, and high-frequency performance (Duclos et al., 1988).

Most research on ICL damping treatments thus far is focused on bending vibration control of isotropic Euler-Bernoulli beams. For example, Agnes and Napolitano (1993) assume a one-mode analysis to approximate the damping performance of a simply-supported ICL beam, and check the results with a finite element method. In the meantime, Baz (1993) and Shen (1993, 1994b) propose different mathematical models to predict bending vibration control of an isotropic Euler-Bernoulli beam with ICL damping treatments. Furthermore, Baz and Ro (1993a) study bending vibration control of isotropic Euler-Bernoulli beams with partial ICL treatments. Also, Van Nostrand et al. (1993) study transient bending vibration of isotropic Euler-Bernoulli beams with ICL through a combination of finite element methods and experiments. In contrast, research on ICL treatments of composite structures is very limited. The only available literature is by Shen (1994a), who analyzes how ICL can reduce bending and twisting vibrations of composite and isotropic plates.

This research has two goals. The first goal is to develop a mathematical model to predict bending, twisting, and axial vibration response of a composite beam with ICL damping treatments. The second goal is to verify the concept of active constrained layer through laboratory experiments and to evaluate the strength and weakness of this new technique.

To achieve the first goal, there are two entirely different approaches to model vibration of an ICL composite beam: mechanics-of-material approach and composite-plate approach. The mechanics-of-material approach starts from a differential beam element. Then combination of kine-
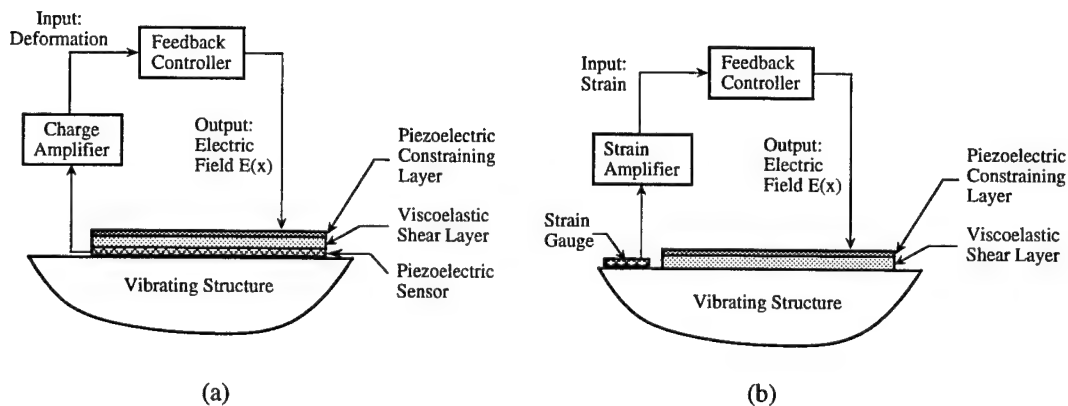
Figure 1: Setup of an active or intelligent constrained layer damping treatment: (a) design by Baz (1993) and (b) design by Shen (1993, 1994b)

matic relation, one-dimensional generalized constitutive equation, and Newton's second law will lead to the equations of motion of ICL composite beams. A potential bottleneck of this approach is that the twisting motion of the ICL composite beam might be too complicated and difficult to model. In contrast, the composite-plate approach considers the composite beam as a composite plate with large length and small width. Based on the results obtained by Shen (1994a) for ICL composite plates, the field equations are integrated with respect to the plate width to yield equations of motion of the composite beam. Though this approach is rigorous, the resulting equations of motion are fairly tedious, because they describe all kinds of possible motion of the composite beam.

In this research, the principal investigator applies the composite-plate approach to obtain the equations of motion of a composite beam. After the field equations of an ICL composite plate is integrated with respect to the plate width, the displacement fields of the plate are linearized along the width. The linearization is valid in low frequency ranges, for which the wave length is greater than the width of the composite beam. As a consequence of the linearization, the response of the composite beam is completely characterized by six quantities: out-of-plane bending deflection, angle of twist, in-plane bending deflection, axial deformation, in-plane deformation along the width, and in-plane shear deformation. Because the resulting equations of motion are very complicated, predicting vibration response through homogeneous solutions and match of boundary conditions becomes impractical. Instead, the equations of motion are normalized and rearranged in a matrix state-space form. Then the transfer function method by Yang and Tan (1992) is applied to predict the coupled bending, torsional, and axial vibration response. Finally, a numerical example is illustrated on a composite beam with bending-twisting coupling stiffness. Because the derivation is tedious and

will by all means exceed the 20-page limit, only important results are summarized in this report. The detailed derivation can be found in the paper by Shen (1994d) submitted to *Smart Materials and Structures*.

To achieve the second goal, two different ICL setups are tested on a composite beam without bending-torsion coupling. According to existing models (Baz, 1993; Shen, 1993, 1994b), a displacement (proportional) feedback controller will increase the damping of ICL because of the damping of the viscoelastic shear layer. Also, a PPF feedback controller will increase the damping of ICL because of the elasticity of the viscoelastic layer. Both theoretical predictions were observed in the experiments and the results agree with those obtained by other researches (Van Nostrand et al., 1993; Baz and Ro, 1994; Azvine et al., 1994)

ICL Composite Beam Theory

This section summarizes some important results of the ICL composite beam theory in the Laplace transform domain. The detailed derivation is given in the paper by Shen (1994d).

Consider a rectangular ICL composite beam with width b (along the x-axis) and length l (along the y-axis) as shown in Fig. 2(a). The composite beam has thickness h_p, density ρ_p, extensional stiffness matrix $\mathbf{A_p}$, bending stiffness matrix $\mathbf{D_p}$, and bending-extension coupling stiffness matrix $\mathbf{B_p}$. The piezoelectric constraining layer has thickness h_c, Young's modulus E_c, Poisson ratio ν_c, piezoelectric constant d_{31}, and density ρ_c. According to the classical plate theory, the corresponding extensional stiffness matrix $\mathbf{A}_c$ and bending stiffness matrix $\mathbf{D}_c$ of the constraining layer are

$$\mathbf{A}_c \equiv \frac{E_c h_c}{1 - \nu_c^2} \begin{bmatrix} 1 & \nu_c & 0 \\ \nu_c & 1 & 0 \\ 0 & 0 & \frac{1-\nu_c}{2} \end{bmatrix} \tag{1}$$

and

$$\mathbf{D}_c \equiv \frac{E_c h_c^3}{12(1 - \nu_c^2)} \begin{bmatrix} 1 & \nu_c & 0 \\ \nu_c & 1 & 0 \\ 0 & 0 & \frac{1-\nu_c}{2} \end{bmatrix} \tag{2}$$

respectively. The shear layer has thickness h_s, density ρ_s, and shear damping modulus $G(s)$, which becomes the shear complex modulus $G(i\omega)$ when the Laplace transform variable s is replaced by $i\omega$. In addition, the shear damping modulus $G(s)$ is related to the shear relaxation function $\chi(t)$ of the viscoelastic layer through

$$G(s) = s \int_0^\infty e^{-st} \chi(t) dt \tag{3}$$
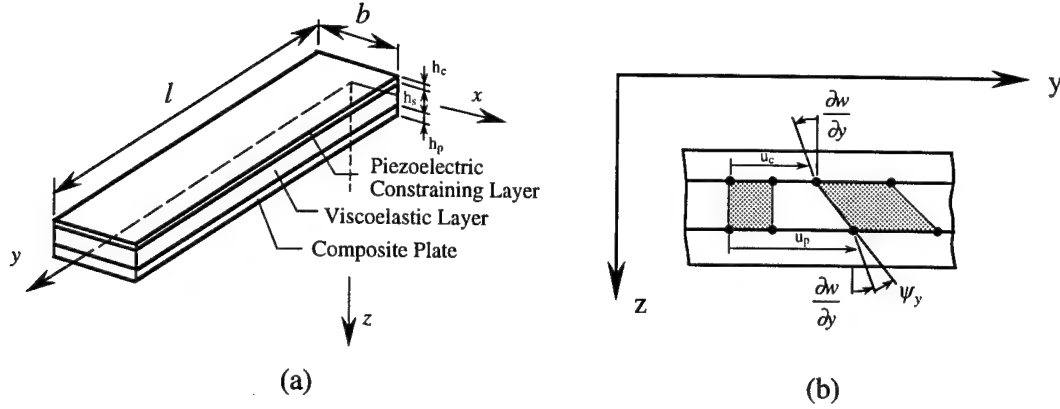
Figure 2: (a) A composite beam with ICL treatments of length l and width b, and (b) deformation of the shear layer in the xz plane

The composite beam is subjected to a laterally applied body force $q_p(x, y, s)$ and an in-plane applied body force $f_p(x, y, s)$. Similarly, the constraining layer is subjected to a laterally applied body force $q_c(x, y, s)$ and an in-plane applied body force $f_c(x, y, s)$. In the sequel, subscripts (and sometimes superscripts) c, s, and p refer to the constraining layer, the viscoelastic shear layer, and the composite beam, respectively. The boundaries at $x = \pm b/2$ are free, and the boundaries at $y = 0$ and $y = l$ can be free, fixed, or simply supported.

Vibration response of the ICL beam shown in Fig. 2(a) is characterized by the deflection $w(x, y, s)$ of the composite beam, the mid-plane displacements $u_p(x, y, s)$ and $v_p(x, y, s)$ of the composite beam in the x and y directions, and the mid-plane displacements $u_c(x, y, s)$ and $v_c(x, y, s)$ of the constraining layer in the x and y directions. In addition, the composite beam, the constraining layer, and the viscoelastic shear layer have the same deflection $w(x, y, s)$. For the composite beam and the constraining layer, the primary deformations are bending and mid-plane deformations. For the viscoelastic layer, the primary deformation is the shear deformation described by the shear angles ψ_x and ψ_y; see Fig. 2(b).

Because the width b is much smaller than the length l of the composite beam, the displacement fields $u_p(x, y)$, $v_p(x, y)$, $u_c(x, y)$, $v_c(x, y)$, and $w(x, y)$ can be approximated by linear distributions in x for low frequency ranges, for which the wavelength is greater than the width b. For example,

$$w(x, y) \approx \bar{w}(y) - x\theta(y) \tag{4}$$

$$u_p(x, y) \approx \bar{u}_p(y) + x\alpha_p(y) \tag{5}$$

31-6

and

$$v_p(x, y) \approx \bar{v}_p(y) + x\beta_p(y) \tag{6}$$

The assumption of linear displacement fields in x provides simple physical interpretation. For example, $\bar{w}(y)$ is the average bending deflection of the ICL beam in the y-z plane, $\bar{u}_p(y)$ is the average in-plane bending deflection of the composite beam in the x-y plane, $\bar{v}_p(y)$ is the average axial motion of the composite beam along the y axis, $\theta(y)$ is the angle of twist of the ICL beam about the y axis, $\alpha_p(y)$ is the in-plane normal strain of the composite beam in the x direction, and β_p is a measure of in-plane shear strain of the composite beam in the x-y plane, as indicated by

$$\gamma_{xy}^{(p)} = \frac{\partial u_p}{\partial y} + \frac{\partial v_p}{\partial x} \approx \frac{\partial \bar{u}_p}{\partial y} + \beta(y) \tag{7}$$

The same displacement fields and phisical meaning can be obtained for the constraining layer as well, if the subscript p in (5) and (6) is replaced by c.

Under the above linearization, the resulting equations of motion need to be normalized to facilitate accurate numerical results. To do so, define the following dimensionless quantities

$$\mathbf{u}_p \equiv \frac{1}{\delta_p} \begin{pmatrix} \bar{u}_p/r \\ \bar{v}_p/r \\ \alpha_p \\ \beta_p \end{pmatrix}, \quad \mathbf{u}_c \equiv \frac{1}{\delta_c} \begin{pmatrix} \bar{u}_c/r \\ \bar{v}_c/r \\ \alpha_c \\ \beta_c \end{pmatrix}, \quad \mathbf{w} \equiv \begin{pmatrix} \bar{w}/r \\ \theta \end{pmatrix}, \quad r \equiv \frac{b}{\sqrt{12}} \tag{8}$$

$$y' \equiv \frac{y}{l}, \quad s' = \frac{s}{p_t}, \quad \eta = \frac{l}{r} \tag{9}$$

In (8) and (9),

$$\delta_p^2 = \frac{D_{22}}{l^2 \sqrt{A_{22}^{(p)} A_{33}^{(p)}}}, \quad \delta_c^2 = \frac{D_{22}}{l^2 \sqrt{A_{22}^{(c)} A_{33}^{(c)}}}, \quad p_t = \sqrt{\frac{D_{22}}{\rho l^4}} \tag{10}$$

where D_{ij}, $A_{ij}^{(p)}$, and $A_{ij}^{(c)}$ are the elements corresponding to the i-th row and the j-th column of $\mathbf{D}_p + \mathbf{D}_c$, $\mathbf{A}_p$, and $\mathbf{A}_c$, respectively. Moreover, define a dimensionless damping

$$\epsilon(s) \equiv \frac{a^2 h_s l^2}{D_{22}} G(s) \tag{11}$$

dimensionless damping coefficients

$$\gamma_p \equiv \frac{l\delta_p}{ah_s}, \quad \gamma_c \equiv \frac{l\delta_c}{ah_s} \tag{12}$$

and dimensionless wave speeds

$$c_p \equiv \sqrt{\frac{\rho_p h_p}{\rho}} \delta_p, \quad c_c \equiv \sqrt{\frac{\rho_c h_c}{\rho}} \delta_c \tag{13}$$

Because the normalized quantities y' and s' in (9) will occur frequently for the rest of the paper, the primes in (9) are dropped; therefore, y and s will refer to the normalized quantities in the sequel.

According to Shen (1994d), the vibration of an ICL composite beam satisfies the following normalized equation of motion.

$$
\left\{
\begin{bmatrix}
-\mathbf{A}_1^{(p)}\frac{\partial^2}{\partial y^2} & \mathbf{A}_2^{(p)}\frac{\partial}{\partial y} & \mathbf{0} & \mathbf{0} & \mathbf{b}_1\frac{\partial^3}{\partial y^3} & \mathbf{b}_2\frac{\partial^2}{\partial y^2} \\
\mathbf{A}_3^{(p)}\frac{\partial}{\partial y} & \mathbf{A}_4^{(p)}-\mathbf{A}_1^{(p)}\frac{\partial^2}{\partial y^2} & \mathbf{0} & \mathbf{0} & \mathbf{b}_3\frac{\partial^2}{\partial y^2} & \mathbf{b}_4\frac{\partial}{\partial y}-\mathbf{b}_1\frac{\partial^3}{\partial y^3} \\
\mathbf{0} & \mathbf{0} & -\mathbf{A}_1^{(c)}\frac{\partial^2}{\partial y^2} & \mathbf{A}_2^{(c)}\frac{\partial}{\partial y} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{A}_3^{(c)}\frac{\partial}{\partial y} & \mathbf{A}_4^{(c)}-\mathbf{A}_1^{(c)}\frac{\partial^2}{\partial y^2} & \mathbf{0} & \mathbf{0} \\
-\mathbf{b}_1^T\frac{\partial^3}{\partial y^3} & \mathbf{b}_3^T\frac{\partial^2}{\partial y^2} & \mathbf{0} & \mathbf{0} & \frac{\partial^4}{\partial y^4} & e_{23}\frac{\partial^3}{\partial y^3} \\
\mathbf{b}_2^T\frac{\partial^2}{\partial y^2} & -\mathbf{b}_4^T\frac{\partial}{\partial y}+\mathbf{b}_1^T\frac{\partial^3}{\partial y^3} & \mathbf{0} & \mathbf{0} & e_{32}\frac{\partial^3}{\partial y^3} & e_{33}\frac{\partial^2}{\partial y^2}+\frac{\partial^4}{\partial y^4}
\end{bmatrix}
\right.
$$

$$
+\epsilon(s)
\begin{bmatrix}
\gamma_p^2 & 0 & 0 & 0 & -\gamma_p\gamma_c & 0 & 0 & 0 & 0 & -\eta\gamma_p \\
0 & \gamma_p^2 & 0 & 0 & 0 & -\gamma_p\gamma_c & 0 & 0 & \gamma_p\frac{\partial}{\partial y} & 0 \\
0 & 0 & \gamma_p^2 & 0 & 0 & 0 & -\gamma_p\gamma_c & 0 & 0 & 0 \\
0 & 0 & 0 & \gamma_p^2 & 0 & 0 & 0 & -\gamma_p\gamma_c & 0 & -\gamma_p\frac{\partial}{\partial y} \\
-\gamma_p\gamma_c & 0 & 0 & 0 & \gamma_c^2 & 0 & 0 & 0 & 0 & \eta\gamma_c \\
0 & -\gamma_p\gamma_c & 0 & 0 & 0 & \gamma_c^2 & 0 & 0 & -\gamma_c\frac{\partial}{\partial y} & 0 \\
0 & 0 & -\gamma_p\gamma_c & 0 & 0 & 0 & \gamma_c^2 & 0 & 0 & 0 \\
0 & 0 & 0 & -\gamma_p\gamma_c & 0 & 0 & 0 & \gamma_c^2 & 0 & \gamma_c\frac{\partial}{\partial y} \\
0 & -\gamma_p\frac{\partial}{\partial y} & 0 & 0 & 0 & \gamma_c\frac{\partial}{\partial y} & 0 & 0 & -\frac{\partial^2}{\partial y^2} & 0 \\
-\eta\gamma_p & 0 & 0 & \gamma_p\frac{\partial}{\partial y} & \eta\gamma_c & 0 & 0 & -\gamma_c\frac{\partial}{\partial y} & 0 & \eta^2-\frac{\partial^2}{\partial y^2}
\end{bmatrix}
$$

$$
+s^2
\begin{bmatrix}
c_p^2\mathbf{I}_4 & 0 & 0 \\
0 & c_c^2\mathbf{I}_4 & 0 \\
0 & 0 & \mathbf{I}_2
\end{bmatrix}
\left.\right\}
\begin{pmatrix} \mathbf{u}_p \\ \mathbf{u}_c \\ \mathbf{w} \end{pmatrix}
= \frac{l^4}{rD_{22}}
\left[
\begin{pmatrix}
\delta_p \bar{f}_x^{(p)} \\
\delta_p \bar{f}_y^{(p)} \\
\delta_p \tilde{f}_x^{(p)}/r \\
\delta_p \tilde{f}_y^{(p)}/r \\
\delta_c \bar{f}_x^{(c)} \\
\delta_c \bar{f}_y^{(c)} \\
\delta_c \tilde{f}_x^{(c)}/r \\
\delta_c \tilde{f}_y^{(c)}/r \\
\bar{q} \\
-\tilde{q}/r
\end{pmatrix}
+ \frac{E_c d_{31}\delta_c}{l(1-\nu_c)}
\begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
0 \\
-\frac{\partial \Delta\bar{V}}{\partial y} \\
\eta\Delta\bar{V} \\
-\frac{1}{r}\frac{\partial \Delta\tilde{V}}{\partial y} \\
0 \\
0
\end{pmatrix}
\right]
\tag{14}
$$

where the symmetry of $\mathbf{B}_p$ has been assumed, and

$$
\mathbf{A}_1^{(p)} \equiv \frac{1}{\sqrt{A_{22}^{(p)}A_{33}^{(p)}}}
\begin{bmatrix} A_{33}^{(p)} & A_{32}^{(p)} \\ A_{23}^{(p)} & A_{22}^{(p)} \end{bmatrix},
\quad
\mathbf{A}_2^{(p)} \equiv -\frac{\eta}{\sqrt{A_{22}^{(p)}A_{33}^{(p)}}}
\begin{bmatrix} A_{31}^{(p)} & A_{33}^{(p)} \\ A_{21}^{(p)} & A_{23}^{(p)} \end{bmatrix}
\tag{15}
$$

31-8

$$\mathbf{A}_3^{(p)} \equiv \frac{\eta}{\sqrt{A_{22}^{(p)} A_{33}^{(p)}}} \begin{bmatrix} A_{13}^{(p)} & A_{12}^{(p)} \\ A_{33}^{(p)} & A_{32}^{(p)} \end{bmatrix}, \quad \mathbf{A}_4^{(p)} \equiv \frac{\eta^2}{\sqrt{A_{22}^{(p)} A_{33}^{(p)}}} \begin{bmatrix} A_{11}^{(p)} & A_{13}^{(p)} \\ A_{31}^{(p)} & A_{33}^{(p)} \end{bmatrix} \tag{16}$$

$$\mathbf{b}_1 = \frac{l\delta_p}{D_{22}} \begin{pmatrix} B_{32} \\ B_{22} \end{pmatrix}, \quad \mathbf{b}_2 = -\frac{2\eta l\delta_p}{D_{22}} \begin{pmatrix} B_{33} \\ B_{23} \end{pmatrix} \tag{17}$$

$$\mathbf{b}_3 = -\frac{\eta l\delta_p}{D_{22}} \begin{pmatrix} B_{12} \\ B_{32} \end{pmatrix}, \quad \mathbf{b}_4 = \frac{2\eta^2 l\delta_p}{D_{22}} \begin{pmatrix} B_{13} \\ B_{33} \end{pmatrix} \tag{18}$$

and

$$e_{23} \equiv -2\eta \frac{D_{23}}{D_{22}}, \ e_{32} \equiv 2\eta \frac{D_{32}}{D_{22}}, \ e_{33} \equiv -4\eta^2 \frac{D_{33}}{D_{22}} \tag{19}$$

In addition, $\mathbf{A}_i^{(c)}$, $i = 1, \ldots, 4$, are defined through (15) and (16) with p replaced by c.

In addition to the equation of motion (14), boundary conditions in terms of $\mathbf{u}_p$, $\mathbf{u}_c$, and $\mathbf{w}$ need to be satisfied at $y = 0$ and $y = l$ as follows.

(i) Restrained to deform laterally: This requires

$$\mathbf{w} = \mathbf{0} \tag{20}$$

(ii) Restrained to rotate: This requires

$$\frac{\partial \mathbf{w}}{\partial y} = 0 \tag{21}$$

(iii) Free of bending moment: This requires

$$\begin{bmatrix} \mathbf{b}_1^T \frac{\partial}{\partial y} & -\mathbf{b}_3^T \\ 0 & \mathbf{b}_1^T \frac{\partial}{\partial y} \end{bmatrix} \mathbf{u}_p + \begin{bmatrix} -\frac{\partial^2}{\partial y^2} & -e_{23} \frac{\partial}{\partial y} \\ 0 & \frac{\partial^2}{\partial y^2} \end{bmatrix} \mathbf{w} = \mathbf{0} \tag{22}$$

(iv) Free of lateral force: This requires

$$\begin{bmatrix} \mathbf{b}_1^T \frac{\partial^2}{\partial y^2} & -\mathbf{b}_3^T \frac{\partial}{\partial y} \\ \mathbf{b}_2^T \frac{\partial}{\partial y} & \mathbf{b}_1^T \frac{\partial^2}{\partial y^2} - \mathbf{b}_4^T \end{bmatrix} \mathbf{u}_p + \begin{bmatrix} -\frac{\partial^3}{\partial y^3} & -e_{23} \frac{\partial^2}{\partial y^2} \\ e_{32} \frac{\partial^2}{\partial y^2} & \frac{\partial^3}{\partial y^3} + e_{33} \frac{\partial}{\partial y} \end{bmatrix} \mathbf{w}$$

$$+ \epsilon(s) \begin{bmatrix} 0 & \gamma_p & 0 & 0 & 0 & -\gamma_c & 0 & 0 & \frac{\partial}{\partial y} & 0 \\ 0 & 0 & 0 & \gamma_p & 0 & 0 & 0 & -\gamma_c & 0 & -\frac{\partial}{\partial y} \end{bmatrix} \begin{pmatrix} \mathbf{u}_p \\ \mathbf{u}_c \\ \mathbf{w} \end{pmatrix} = \mathbf{0} \tag{23}$$

(v) Restrained to deform in the in-plane direction: This requires

$$\mathbf{u}_p = \mathbf{0} \tag{24}$$

for the composite beam and

$$\mathbf{u}_c = 0 \tag{25}$$

for the constraining layer.

(vi) Free of mid-plane laoding: This requires

$$\left[\begin{array}{cc} \mathbf{A}_1^{(p)} \frac{\partial}{\partial y} & -\mathbf{A}_2^{(p)} \\ 0 & \mathbf{A}_1^{(p)} \frac{\partial}{\partial y} \end{array} \right] \mathbf{u}_p + \left[\begin{array}{cc} -\mathbf{b}_1 \frac{\partial^2}{\partial y^2} & -\mathbf{b}_2 \frac{\partial}{\partial y} \\ 0 & \mathbf{b}_1 \frac{\partial^2}{\partial y^2} \end{array} \right] \mathbf{w} = \mathbf{0} \tag{26}$$

for the composite beam and

$$\left[\begin{array}{cc} \mathbf{A}_1^{(c)} \frac{\partial}{\partial y} & -\mathbf{A}_2^{(c)} \\ 0 & \mathbf{A}_1^{(c)} \frac{\partial}{\partial y} \end{array} \right] \mathbf{u}_c - \frac{E_c d_{31} l}{(1 - \nu_c) r \delta_c \sqrt{A_{22}^{(c)} A_{33}^{(c)}}} \left(\begin{array}{c} 0 \\ \Delta \bar{V} \\ 0 \\ \Delta \tilde{V} \end{array} \right) = \mathbf{0} \tag{27}$$

for the constraining layer.

A Numerical Example

Consider a composite beam of 2 cm wide and 10 cm long consisting of $[45°/-45°/-45°/45°]$ angle ply laminate. Each lamina is 0.6-mm thick and contains 60 volume percent of T-300 carbon fiber in an epoxy matrix. The stiffness matrices are

$$\mathbf{A} = \left[\begin{array}{ccc} 962.64 & 806.64 & 0 \\ 806.64 & 962.64 & 0 \\ 0 & 0 & 829.68 \end{array} \right] \times 10^5 \frac{\mathrm{N}}{\mathrm{m}} \tag{28}$$

$$\mathbf{D} = \left[\begin{array}{ccc} 46.21 & 38.72 & 27.04 \\ 38.72 & 46.21 & 27.04 \\ 27.04 & 27.04 & 39.82 \end{array} \right] \mathrm{N} \tag{29}$$

and $\mathbf{B} = 0$ because the composite laminate is symmetric.

In addition, the specific gravities of T-300 and the epoxy are estimated to be 1.76 and 1.25, respectively. Therefore, the specific gravity of the composite will be $1.76(0.6) + 1.25(0.4) = 1.556$.

The piezoelectric cover sheet is PZT-5Z with thickness 0.38 mm, density 7,500 kg m^{-3}, Young's modulus 49 Gpa, and Poisson ratio 0.3 (estimated). The viscoelastic shear layer has thickness of 1 mm and the complex modulus is assumed to be

$$G(\omega) = 0.142 \left(\frac{\omega}{2\pi} \right)^{0.494} (1 + 1.46i)(\mathrm{MNm}^{-2}) \tag{30}$$
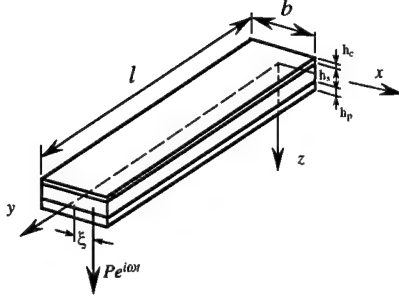
Figure 3: A cantilever ICL composite beam subjected to an offset concentrated load

The composite beam is cantilevered; fixed at $y = 0$ and subjected to a concentrated harmonic load $Pe^{i\omega t}$ at the free end with offset ξ (see Fig. 3). In addition, the deflection at the free end (i.e., $y = 1$ after normalization) is fed back to actuate a uniform voltage ΔV across the thickness of the constraining layer through an arbitrary dimensionless transfer function $C(s)$, i.e.,

$$\Delta V = -\frac{r(1 - \nu_c)\delta_c \sqrt{A_{22}^{(c)} A_{33}^{(c)}}}{E_c d_{31} l} \lambda C(s) w(1, s) \tag{31}$$

where λ is a dimensionless control gain and $w(1, s)$ is the dimensionless deflection at the free end. Because ΔV is uniform, $\Delta \bar{V}$ is a constant and $\Delta \tilde{V} = 0$. Moreover, assume that a reactive force $f_x^{(c)}$ exists such that $\mathbf{f} = 0$ in (14). The goal of this example is to study how the control gain λ of the ICL will affect bending and torsional vibration of the composite beam.

Because the equation of motion and boundary conditions of this system are so complicated, it is no longer practical to find the homogeneous solution of (14) and to determine its coefficients through boundary conditions. To obtain a tangible solution of (14), the transfer function approach by Yang and Tan (1992) is used. First of all, define a state vector

$$\mathbf{x}(y, s) \equiv \left(\mathbf{u}_p^T, \mathbf{u}_c^T, \mathbf{w}^T, \frac{\partial^2 \mathbf{w}^T}{\partial y^2}, \frac{\partial \mathbf{u}_p^T}{\partial y}, \frac{\partial \mathbf{u}_c^T}{\partial y}, \frac{\partial \mathbf{w}^T}{\partial y}, \frac{\partial^3 \mathbf{w}^T}{\partial y^3} \right)^T \tag{32}$$

Therefore, the equation of motion (14) and the corresponding boundary conditions can be rewritten as

$$\mathbf{F}_1(s)\frac{\partial \mathbf{x}(y, s)}{\partial y} + \mathbf{F}_2(s)\mathbf{x}(y, s) = 0 \tag{33}$$

and

$$\mathbf{M}(s)\mathbf{x}(0, s) + \mathbf{N}(s)\mathbf{x}(1, s) = \mathbf{g}(s) \tag{34}$$

where $s = i\omega$ and explicit expressions of $\mathbf{F}_1(s)$, $\mathbf{F}_2(s)$, $\mathbf{M}(s)$, and $\mathbf{N}(s)$ can be found in Shen (1994d). Accordng to Yang and Tan (1992), the response is obtained by

$$\mathbf{x}(y) = e^{-\mathbf{F}(s)(1-y)} \left[\mathbf{M}e^{-\mathbf{F}(s)} + \mathbf{N} \right]^{-1} \mathbf{g}(s) \tag{35}$$

where

$$\mathbf{F}(s) \equiv \mathbf{F}_1^{-1}(s)\mathbf{F}_2(s) \tag{36}$$
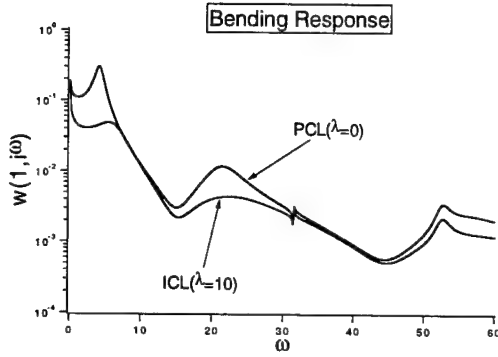
Figure 4: Bending response $w(1, i\omega)$ at the free end
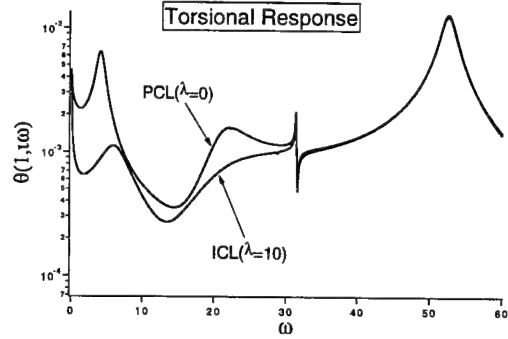
Figure 5: Torsional response $\theta(1, i\omega)$ at the free end

As a numerical example, let $\xi = 1$ cm and the controller is a proportional-plus-derivative controller with

$$C(s) = 1 + 0.1s \tag{37}$$

Figures 4 and 5 show the bending vibration response $w(1, i\omega)$ and torsional vibration response $\theta(1, i\omega)$ at the free end. Before the feedback control is activated, the ICL damping behaves like a traditional passive constrained layer. The system has four vibration modes for $0 < \omega < 60$. The bending response of all four vibration modes are reduced when the control is applied. As far as the torsional response, however, only the first two modes are reduced simultaneously by the ICL bending damper. (The reduction of vibration response is shown in Figs. 6 and 7.) This phenomenon is a direct consequence of the control law (31). When the control is not activated, Fig. 4 shows that the third and fourth modes have much smaller bending response than the first and second modes. According to (31), the third and fourth modes receive smaller control actions resulting in smaller reduction in coupled bending and torsional vibration. This can also be shown through a sensitivity analysis as follows. Let $y = 1$ in (35) to obtain

$$\left[\mathbf{M}e^{-\mathbf{F}} + \mathbf{N}(\lambda) \right] \mathbf{x}(1, \lambda) = \mathbf{g} \tag{38}$$

Notice that the boundary condition matrix $\mathbf{N}$ and the state vector $\mathbf{x}$ both depend on the control gain λ. Differentiate (38) with respect to λ to obtain

$$\left[\mathbf{M}e^{-\mathbf{F}} + \mathbf{N}(\lambda) \right] \frac{d\mathbf{x}(1, \lambda)}{d\lambda} + \frac{d\mathbf{N}(\lambda)}{d\lambda} \mathbf{x}(1, \lambda) = \mathbf{0} \tag{39}$$

which can be simplified to (Shen, 1994d)

$$\frac{d\mathbf{x}(1, \lambda)}{d\lambda} = -C(s)w(1, \lambda)\mathbf{v}_{18} \tag{40}$$

31-12

Mode 1

Mode 2
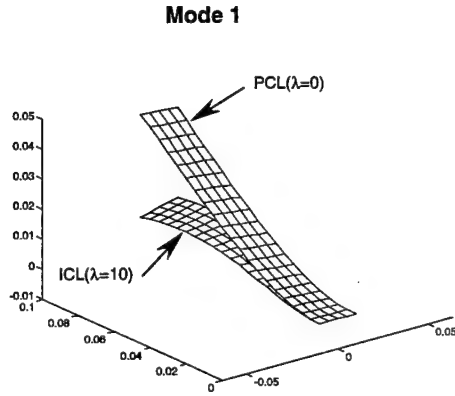
Figure 6: The first vibration mode at $\omega = 4.0$ before and at $\omega = 5.4$ after the control is applied
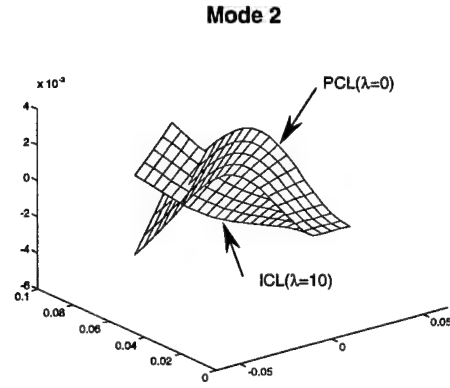
Figure 7: The second vibration mode at $\omega = 21.8$ before and at $\omega = 22.4$ after the control is applied

where $\mathbf{v}_{18}$ is the 18th column of $\left[\mathbf{M}e^{-\mathbf{F}(s)} + \mathbf{N}(\lambda)\right]^{-1}$. From (40), the control gain λ will not change the state vector $\mathbf{x}$ significantly when $w(1)$ is small.

From the sensitivity analysis above, it is natural to propose that the angle of twist $\theta(1, s)$ be used instead of $w(1, s)$ in the control law (31) to control the third and fourth modes where torsional vibration is more significant than the bending vibration. This proposal, however, could be very misleading. For example, if the m-th mode shape has positive bending and positive twisting, and the n-th mode shape has positive bending and negative twisting, then using $\theta(1, s)$ to control bending vibration will reduce the vibration of one of the mode but increase the vibration of the other mode. A possible way to improve the control efficiency might be to use a nonlinear control law such as

$$\Delta V = -C(s)w(1, s)\left|\theta(1, s)\right| \tag{41}$$

Experimental Evaluations

Figure 8 shows two different setups used in the experiments to prove the concept of active constrained layer. Setup (a), used by Azvine et al. (1994), consists of a PZT actuator 2 mounted directly on to the 3M passive constrained layer. In addition, a PZT acutator 1 collocated with a PZT sensor is attached on the other side of the composite beam. In constrast, setup (b), used by Van Nostrand et al. (1993), consists of three partially treated active constrained layers. The PZT actuator 2 is glued directly to the 3M ISD112 viscoelastic material. Setup (b) also has a sensor and the actuator 1 on the
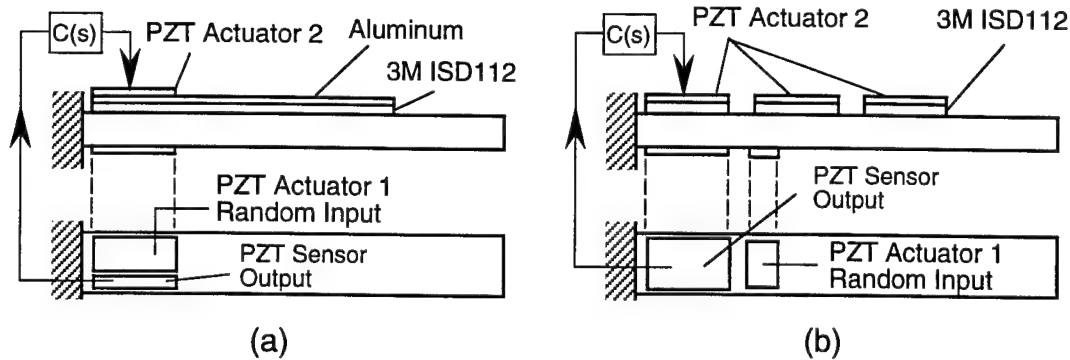
Figure 8: Two different setups used in the active constrained layer experiments

other side of the beam. In addition, the sensor is collocated with the actuator 2. Notice that setup (a) has an additional aluminum constraining layer; therefore, the constrained layer treatment occupies a large area of the beam. On the contrary, setup (b) uses PZT actuator 2 as the constraining layer; therefore, the damping treatment occupies a smaller area of the beam, because large area of PZT is difficult to find. In both setups, a random disturbance is input to the actuator 1. According to the response from the sensor, the actuator 2 is regulated through a PPF or displacement (proportional) controller. In the experiments, frequency response functions from the actuator 1 to the sensor is measured. The dimensions of the composite beam is 13 in.×1.5 in.×0.05 in. The dimensions of the PZT is 2.5 in.×1.5 in.×0.01 in. The thickness of both the aluminum and the ISD112 is 0.01 in.

According to existing theories (Baz, 1993; Shen, 1993; 1994b), a displacement (proportional) controller for ψ_y in Fig. 2(b) will increase the damping, because the shear deformation ψ_y is increased in the damping layer. In addition, a PPF controller for ψ_y in Fig. 2(b) will increase the damping as well, because the control action is transmitted to the beam through the elasticity of the viscoelastic layer. The primary goals of the experiments are to verify the existence of these two damping mechanisms, and to determine the dimensions and material properties to maximize damping performance of the ACL.

Many experiments are conducted. The following figures summarize the experimental results. Figure 9 shows comparison of the open-loop frequency response and the closed-loop response with a PPF controller of gain 5 for the setup (a). (Figure 10 shows the frequency response of the PPF controller.) The open-loop system corresponds to the traditional passive constrained layer damping treatments. When the PPF controller is activated, the amplitude of the first bending mode at 18 Hz reduces from 264×10^{-3} to 145×10^{-3}. The higher modes are not affected by the PPF controller.

31-14
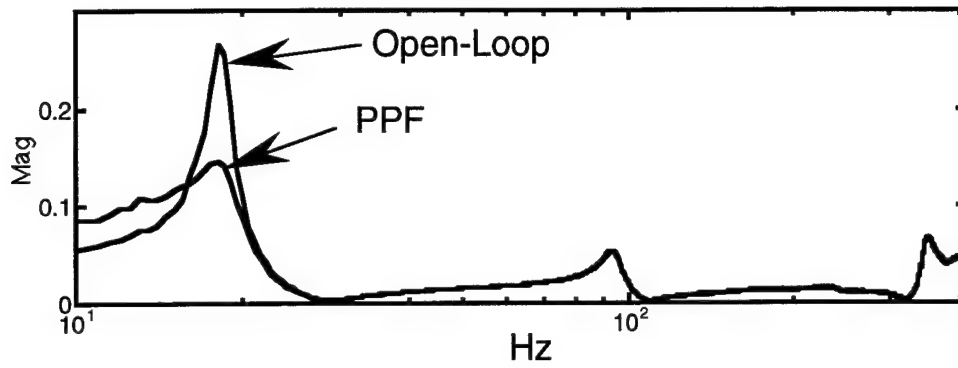
Figure 9: Frequency response functions of setup (a): open-loop without control and closed-loop with PPF controller; gain = 5
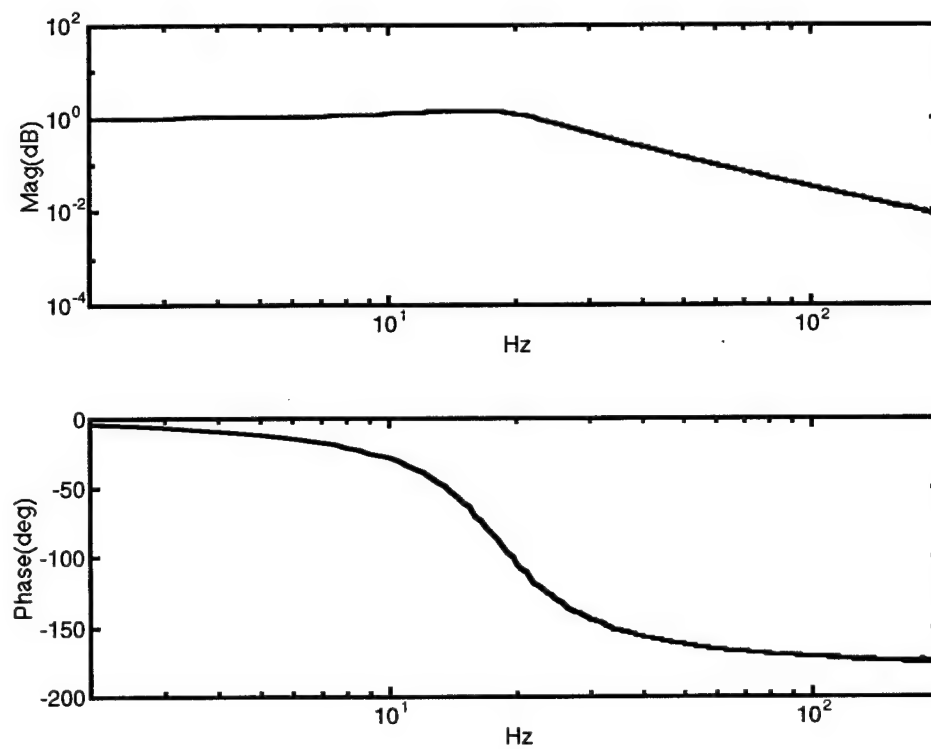
Figure 10: Frequency response of the PPF

31-15
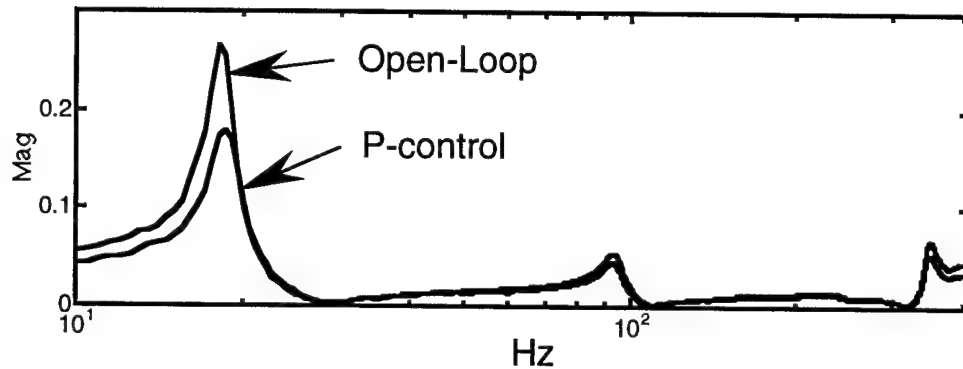
Figure 11: Frequency response functions for setup (a): open-loop without control and closed-loop with proportional (displacement) controller; gain = 5

Figure 11 shows comparison of the open-loop frequency response and the closed-loop response with a proportional (displacement) controller of gain 5 for the setup (a). Notice the the proportional controller reduces the amplitude of the first bending mode from 264×10^{-3} to 173×10^{-3}. Though the proportional controller is not as effective as the PPF controller for the first mode, the proportional controller does reduce the amplitudes of the higher vibration modes a little bit.

Figure 12 shows comparison of the open-loop frequency response and the closed-loop response with proportional (displacement) controller of gain 3 for the setup (b). Notice that the the proportional controller reduces the amplitude of the first bending mode from 136×10^{-3} to 98×10^{-3}. Because the area of the PZT is too small, the control action is relatively limited. In addition, the PZT actuator was found to have a crack around the lead after the experiment, which explains the noisy signal shown in Fig. 12. The experiment would have been repeated if there had been enough time.

Conclusions and Discussions

1. Results of this research lead to an analytical model of ICL composite beams obtained by integrating the ICL plate model with respect to the plate width. The proposed ICL composite beam model is capable of predicting the coupling between bending, axial, and torsional vibrations. The detailed derivation is in a paper submitted to *Smart Materials and Structures* (Shen 1994d).
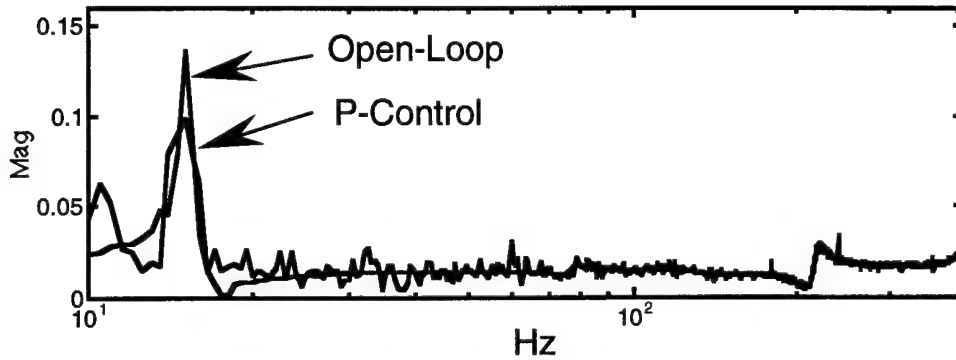
Figure 12: Frequency response functions of the setup (b): open-loop without control and closed-loop with proportional (displacement) controller; gain = 3

2. Numerical results show that ICL bending dampers may or may not reduce coupled bending and torsional vibration of a composite beam simultaneously. When the bending deflection is fed back to activate the ICL bending damper, the sensitivity analysis shows that only those modes with significant bending response will be suppressed simultaneously with their corresponding torsional components. More advanced control theory might be needed in order to make the control more efficient for those modes without significant bending response.

3. The normalization is critical to obtaining accurate numerical results from the transfer function method by Yang and Tan (1992). If the equation of motion is not normalized, ill-conditioned matrices occur and the results might be inaccurate.

4. When the bending-stretching coupling stiffness $\mathbf{B}_p$ is present, the equation of motion (14) remains valid. Because the presence of $\mathbf{B}_p$ increases the order of derivative of the equation of motion, additional boundary conditions are needed to determine the vibration response.

5. The stability of the system can be analyzed through the approach proposed by Shen (1994c).

Experimental observations and existing theory of ACL lead to the following conclusons:

1. An optimization strategy is needed to determine the material properties, thickness, area, and weight penalty of the viscoelastic layer that maximize the damping performance

2. In order to obtain large active damping contribution in ACL treatments, viscoelastic shear layers of high stiffness and small loss factor should be used. These requirements, however, will

31-17

reduce the passive damping from the constrained layer. Therefore, the trade-off between the passive and active damping needs to be considered *apriori*.

3. If the PZT actuator of setup (b) is replaced by PVDF, then the active constrained layer can cover large area. The low stiffness of PVDF, however, might lead to poor performance of the active constrained layer for beams with comparably larger bending stiffness.

4. PPF is known as an effective feedback controller for active damping. Suggested by the experimental results, design of a more effective feedback controller to maximize ACL damping needs to be studied.

Acknowledgements

Bibliography

Agnes G S and Napolitano K 1993 Active constrained layer viscoelastic damping *Proc. 34 SDM Conf.* pp 3499-3506

Azvine B, Wynne R J, and Tomlinson G R 1994 Active damping for the control of flexible structures *CONTROL '94*, March, pp. 1296-1304.

Baz A 1993 Active constrained layer damping *Proc. Damping 93 (San Francisco, CA, 1993)* vol 3 pp IBB1-23

Baz A and Ro J 1993 Partial treatments of flexible beams with active constrained layer damping *Recent Developments in Stability, Vibration and Control of Structural Systems* (ASME AMD vol 167) pp 61-80

Baz A and Ro J 1994 The concept and performance of active constrained layer damping treatments *Sound and Vibration*, March, pp. 18-21.

Duclos T G, Coulter J P and Miller L R Application for smart materials in the field of vibration control *ARO Smart Materials, Structures, and Mathematical Issues Workshop Proceedings*, Virginia Polytechnic Institute and State University, Blacksburg, VA., September 15-16, pp. 132-146.

Shen I Y 1993 Intelligent constrained layers: an innovative approach *Intelligent Structures, Materials and vibrations* (ASME DE vol 58) pp 75-82

Shen I Y 1994a Bending-vibration control of composite and isotropic plates through intelligent constrained-layer treatments *Smart Mater. Struct.* vol 3 pp 59-70

Shen I Y 1994b Hybrid damping through intelligent constrained layer treatments *ASME J. Vib. Acoustics* Vol 116 pp 341-349.

Shen I Y 1994c Stability and controllability of Euler-Bernoulli beams with intelligent constrained layer damping treatments. *ASME J. Vib. Acoustics* in press

Shen I Y 1994d Bending and torsional vibration control of composite beams through intelligent constrained layer damping treatments *Smart Mater. Struct.* under review

Van Nostrand W C, Knowles G J and Inman D J 1993 Active constrained layer damping for microsatellites *Dynamics and control of Structures in Space* vol II ed C L Kirk and P C Hughes pp 667-81

Yang B and Tan C A 1992 Transfer functions of one-dimensional distributed parameter systems *ASME J. Appl Mech.* vol 59 1009-14

Frequency Domain Analysis Of Short Exposure, Photon-Limited Astronomical Data

Melani Shoemaker
Instructor
Department of Electrical Engineering

Seattle Pacific University
Seattle, Washington 98119

Frequency Domain Analysis Of Short Exposure, Photon-Limited Astronomical Data

Melani Shoemaker
Instructor
Department of Electrical Engineering
Seattle Pacific University

Abstract

Obtaining high resolution images of objects in space is a common subject in research today. Turbulence in the atmosphere limits the image resolution obtained with larger telescopes. To combat this atmospheric effect, speckle imaging techniques have been developed that average out atmospheric turbulence to improve image resolution. The performance of these techniques is limited by the achievable signal-to-noise ratios (SNR) in the data. Our aim is to use experimental data to test theories on the effect of spectral bandwidths on the data SNR. In addition, we use experimental data to determine the Fourier domain noise characteristics of the processed data. These results will be used to determine the usefulness of prior knowledge for noise reduction.

FREQUENCY DOMAIN ANALYSIS OF SHORT EXPOSURE, PHOTON-LIMITED ASTRONOMICAL DATA

Melani Shoemaker

1. Introduction

Atmospheric turbulence imposes severe resolution limitations on ground-based telescopes observing space objects over long time exposures (> 60ms). This turbulence results in an effect termed atmospheric seeing in which only the low spatial frequency information passes through the atmosphere free from significant attenuation. Speckle imaging techniques[1,2] have been developed which average out this turbulence to improve image resolution. The amount of resolution improvement is a function of the achievable data signal-to-noise ratio (SNR). Typically, the image data is collected over a narrow spectral bandwidth in order to minimize smearing due to decorrelation of image data as a function of wavelength. However, recently it has been shown[3] theoretically that the data SNR is an increasing function of the spectral bandwidth. Therefore, this theory predicts that the data should be collected with a broad spectral bandwidth instead of a narrow bandwidth. This has not yet been verified experimentally.

In addition, prior knowledge such as positivity and support has long been used to improve the quality of images beyond that achievable without using prior knowledge. Recently, it has been shown[4] that the ratio of the variance of the imaginary component of the Fourier spectrum to that of the real component provides insight into the usefulness of prior knowledge to enhance resolution. For speckle imaging, the processing methods are nonlinear which makes it extremely difficult to analytically calculate the variances of the real and imaginary parts of the reconstructed Fourier spectrum. However, recent computer simulation results at the Phillips Laboratory show large differences in the magnitudes of the variances, which indicates that prior knowledge should be useful for improving image resolution. However, these results have not been validated with field data.

The research this summer had two primary goals. The first goal was to use field data to determine the validity of the theory which predicts that the data SNR is an increasing function of bandwidth. The second goal was to use field data to verify the computer simulation results

concerning the variance properties of Fourier spectra reconstructed using speckle imaging techniques.

The general concept of speckle imaging reconstruction techniques exploits the relationship that the image is the convolution of the true object with the system point spread function. So, to find the true object we estimate the system point spread function as the image obtained from an unresolved star, which acts as a delta function. The object Fourier spectrum is then the Fourier spectrum of the image divided by the Fourier spectrum of the estimated point spread function. Ideally, then, without noise, the object image is simply the inverse Fourier transform of the object Fourier spectrum. Our focus is on image reconstruction of stars at low light levels; as a result, attempts to reconstruct an image from such data require employment of statistical processes to account for atmospheric turbulence and photon noise. Our processes of choice involve power spectrum and bispectrum phase estimation techniques. Image reconstruction in this environment places stringent demands on the characteristics of the data used. Focus, atmospheric seeing, tracking, exposure time, quantity of data, and brightness of the star all have significant effects on the data characteristics that permit good image reconstruction. Therefore, before embarking on reconstructing an image from the data, we reduce and analyze the data for such characteristics.

My work this summer, thus, fell into two categories. One effort has been reduction of various image data sets in an attempt to find data that met the requirements of the reconstruction techniques. Once appropriate data was found, we employed speckle imaging techniques to analyze the data for sufficient power spectrum SNR. If the data was still deemed useful, then the phase was estimated via the bispectrum of the image. The phase and Fourier spectrum estimates reveal further information regarding the quality of the data in question as will be discussed later.

The next section of this paper will briefly describe the speckle imaging techniques applied to the data, as well as the various data formats that were used. Subsequently, the results of the speckle image post processing will be presented for each of three groups of data. Two of the three groups of data were taken with a fixed spectral bandpass and will be useful for our analysis of their Fourier spectrum variances, but not for the spectral bandpass SNR theories. The other data group was taken with varying spectral bandpasses, but we will show that, unfortunately, it

was deemed inadequate for our purposes. Finally, we present a summary of our conclusions and briefly mention our future plans for expanding on the research performed this summer.

2. Methodology

speckle imaging: In the 1970's, Labeyrie[1] discovered that the power spectrum of a short exposure of an atmospherically distorted image retains high spatial frequency information (up to the diffraction limit), although greatly attenuated. Thus, the SNR of the power spectrum at mid to high frequencies will be low. Averaging the power spectrum over many short exposure frames increases this SNR by a factor equal to at least the square root of the number of frames used in the calculation.

Unlike the power spectrum, the phase spectrum is uniformly distributed at spatial frequencies beyond the seeing limit. Therefore methods other than direct averaging must be used in order to obtain phase information beyond this limit. Some methods exist that allow the phase to be indirectly calculated from an intermediate process that does allow many frames to be averaged in order to improve the Fourier spectrum SNR. Our method of choice is based on calculation of the bispectrum of the image, the third moment of the image's Fourier transform, which conveniently retains phase information at high frequencies[2].

power spectrum: Based on Labeyrie's proposals, we obtained a series of short exposure ($\leq$30ms time duration) pictures (frames) to be used in calculating the average power spectrum as follows. To reduce the data into useful form, a biased power spectrum is calculated from each frame as the square of the modulus of the FFT (I(u)) of the short exposure image. To unbias the power spectrum, the estimated frame noise bias (fnoise) is subtracted. Running totals are calculated of the now unbiased power spectra, as well as of the square of the unbiased power spectra. The mean power spectrum (ps) for the entire set of data is the unbiased power spectrum running total, divided by the number of frames (frames) in the data set (1).

$$ps(u) \; = \; \sum_{frames} \left(\left| I(u) \right|^2 - fnoise \right) / \, frames \tag{1}$$

If the data is of sufficient quality, then a low frequency seeing spike should be evident *and* a speckle imaging wing (a plateau of lesser magnitude than the seeing spike, and significant frequency duration) should appear indicating the presence of mid and high spatial frequencies as

extracted from the short exposure images. Next, the power spectrum variance is computed as the square of the mean power spectrum subtracted from the running total of the square of the unbiased power spectra divided by the number of frames (2).

$$\text{var}(u) = \left[\sum_{\text{frames}} \left(|I(u)|^2 - \text{fnoise} \right)^2 / \text{frames} \right] - \text{ps}(u)^2 \tag{2}$$

Finally, the average power spectrum SNR per short exposure frame can be formed as the power spectrum mean divided by the square root of the variance (3).

$$\text{SNR}(u)_{\text{singleframe}} = \text{ps}(u) \Big/ \sqrt{\text{var}(u)} \tag{3}$$

As stated earlier, multiplying by the square root of the number of frames yields the power spectrum SNR for the entire data set (4).

$$\text{SNR}(u)_{\text{entireset}} = \left(\text{ps}(u) \Big/ \sqrt{\text{var}(u)} \right) * \sqrt{\text{frames}} \tag{4}$$

To be deemed useful data, the SNR must be larger than one up to the desired spatial frequency.

bispectrum and phase: In order to estimate phase we calculate the bispectrum of each frame, compute the average bispectrum over all frames, and then calculate the phase from the bispectrum. The bispectrum of a frame, B(u,v), is a product of the Fourier transform of the image frame, $I(u)=I(u_1,u_2)$, as described by equation (5) where * implies the complex conjugate[2].

$$B(u, v) = I(u)I(v)I * (u + v) \tag{5}$$

The phase of the image at frequency (u+v) can therefore be found from the average bispectrum as:

$$\text{ph}\{I * (u + v)\} = \text{ph}\{B(u, v) / [I(u)I(v)]\} \tag{6}$$

Therefore, if B(u,v), I(u), and I(v) are known, then the phase at (u+v) can be calculated. This iterative process requires that two low frequency phases be known. Typically, the phase for I(1,0) and I(0,1) are arbitrarily selected. It is known that the $\text{ph}\{I(0,0)\} = 0$ for a real image. Notice that the bispectrum is four-dimensional in u and v, [$B(u_1,u_2,v_1,v_2)$], creating a huge matrix for a 128x128 image. Fortunately, the bispectrum holds eight-fold symmetry, *and* only a fraction of the non-redundant elements contribute to frequencies with high SNR, allowing image

reconstruction with only calculating a small percentage of the bispectrum. The recursive phase estimation process employed in this research is explained more fully in Matson[5]. The average phase spectrum can then be combined with the previously calculated power spectrum to create the average Fourier spectrum of the image.

prior knowledge theory: Having obtained the Fourier spectrum, the variances of the real and imaginary components can be calculated. Matson[4] has found a link between the ratio of the variances of the Fourier spectrum of an image and the ability to improve image reconstruction based on prior knowledge of the object being imaged. He theorizes that if the variance of the imaginary component of the Fourier spectrum remains either much larger, or much smaller than its real component (more specifically a multiplicative factor of 0.6), then a priori knowledge would be helpful in reducing noise in the image domain.

the cameras: We carried out the above methods on image data of unresolved stars taken from two different camera types that specialize in photon-limited (low light level) images: the Multiple Anode Microchannel Array (MAMA) and Charge Coupled Device (CCD) cameras. The photon-counting MAMA camera provides temporal and spatial coordinates (t, x, y) for each photon event encountered at rates up to 1 MHz. Part of my work this summer involved writing code to reformat the (t, x, y) sequential data into short exposure frames for a given exposure time. The CCD camera, on the other hand, delivers data in frames with an exposure time that is fixed when the data is originally stored. Furthermore, the CCD camera itself generates artifacts due to varying gains and additive bias among pixel locations, which must be quantified and accounted for. This is done by collecting frames of data taken of a uniform ("flat") field and then averaging the image frames to find the multiplicative gain for each pixel. In addition, data is collected in the dark, multiplied by the average flat field, and then averaged in the image domain to yield the average additive bias for each pixel. For data collected by the MAMA camera, the estimated frame noise to use in unbiasing the power spectrum is simply the number of photon events in the frame. For CCD data, the estimated frame noise includes the average read noise induced by the camera itself added to the sum of the number of photon events in a frame.
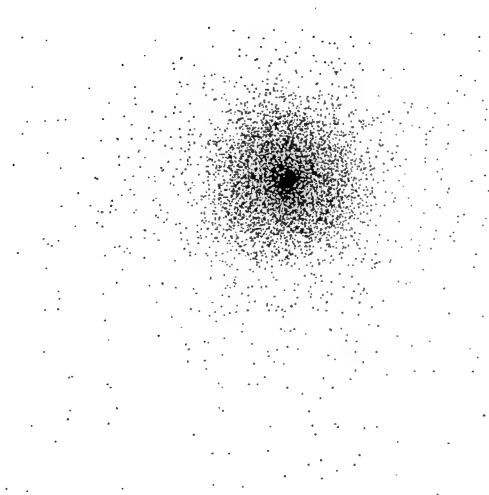
Figure 1 Short exposure of SAO 68166
9.6ms time duration short exposure image of reference star
SAO 68166 taken at the Starfire Optical Range (SOR) at
Kirtland Air Force Base, June 1993.

3. Results

The three data sources to which we applied speckle imaging include: data provided by
Keith Hege of Steward Observatory taken with a MAMA camera in October 1986; data taken
with a MAMA camera at the Starfire Optical Range (SOR) of Phillips Laboratory at Kirtland
AFB, June 1993; and CCD camera data taken at the Air Force Maui Optical Site (AMOS) in
March of 1992. The data taken at the SOR is the group that includes varied spectral bandpasses.

For reference, Figure 1 shows a short exposure image taken at the SOR with a 9.6ms time
exposure of the star SAO 68166 of magnitude 7.0 (G5). As seen in the picture, photon-limited
short exposure images (~10-30 ms) will be sparse: for the data presented in this paper, the
photon events per short exposure frame are in the range from 350 to 7000. The image frame of
Figure 1 is typical of the short exposure data used in this paper (although, this is the brightest
with 7000 photons/frame).

Hege MAMA Data: The data provided by Hege is composed of frames with 10ms time
duration and events/frame on the order of 350. The telescope diameter (D) was 2.3m with a
central wavelength (λ) of 550nm and a fixed spectral bandpass of 30nm. All results for this
group of data are plotted as a function of spatial frequency normalized to one at the diffraction
limit (D/λ). Of the various sets available, we processed images of 66 Orionis, 130 Tauri,
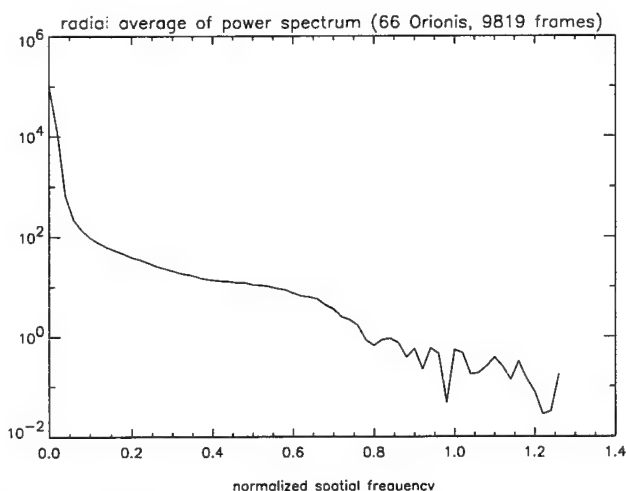
Figure 2 Radial average power spectrum of 66 Orionis
Radial average of the mean power spectrum of the star 66
Orionis taken at Steward Observatory in October 1986.
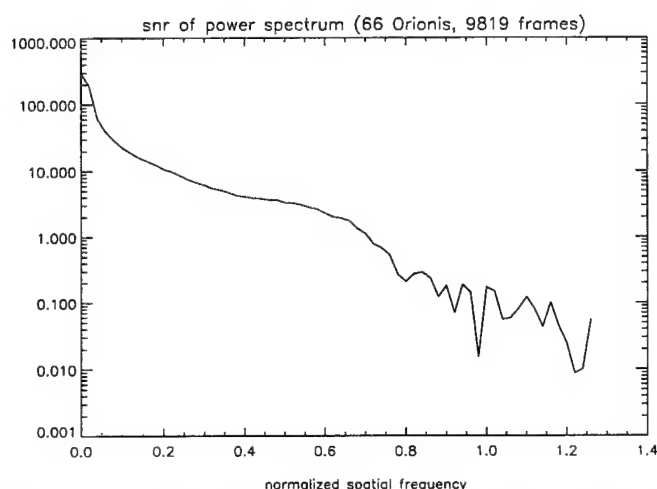
Figure 3 Power spectrum signal-to-noise ratio (SNR) of 66 Orionis
SNR of the radial average power spectrum of the star 66
Orionis taken at Steward Observatory in October 1986.

and Zeta Delphinus. Since all three sets yielded similar results, only the specifics of the 66

Orionis data are presented here. To reiterate, our interest in this data is in analyzing the image's

Fourier spectrum variances. To accomplish this, we first calculated the average power spectrum.

Figure 2 shows the radial average power spectrum for 66 Orionis over 9819 10ms frames with an

average of 313 events/frame. This figure illustrates the typical, expected average power spectrum

for all photon-limited stars: seeing spike at very low frequencies (0-0.03); further high frequency

information present, but attenuated (0.03-0.75); and frequencies near the diffraction limit (1.0)

are clouded by noise (0.8+). For quality data, the mid frequency speckle imaging wing should exist as seen in Figure 2. The radial average power spectrum SNR for the entire data set is shown in Figure 3. Note that the SNR remains greater than one up to approximately 0.75, implying that data up to well over half the diffraction limit has strong signal information.

So far, this group of data meets our requirements and appears likely to be useful in the Fourier spectrum analysis. Next, we calculated the bispectrum and subsequently the phase and Fourier spectrum for each data set as described above using a cutoff frequency equal to the diffraction limit. The estimated phase for 66 Orionis is provided in Figure 4. In agreement with the power spectrum, the phase plot demonstrates strong signal information from d.c. (the center of the circle) outward to well over half the diffraction limit (just prior to the edge of the circle). This is evident by the smooth nature of the interior of the phase plot, and the rough outer edges. [Note, the phase of Figure 4 appears to have discontinuities in the signal; however, the sharp visual contrast between black and white is actually a smooth transition resulting from the black representing a phase of $-\pi$ and the white, $+\pi$.]

Earlier it was mentioned that noise reduction with applied prior knowledge is possible if the Fourier spectrum variances are disparate by at least a factor of 0.6. This group of data demonstrates this phenomenon: Figure 5. Note in this figure the ratio of the variance of the imaginary component to the real is less than 0.6 for the frequencies of the seeing spike (0-0.3), and that the ratio is greater than 1/0.6 for mid-range frequencies (until 0.2).
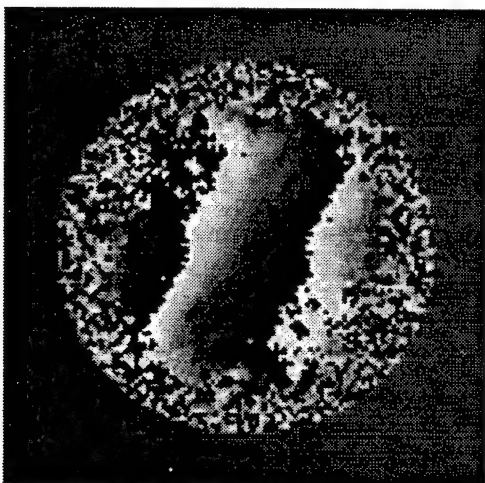
Figure 4 Estimated Phase of the 66 Orionis image
Picture of the estimated phase of the star 66 Orionis. Black corresponds to $-\pi$ and white to $+\pi$.
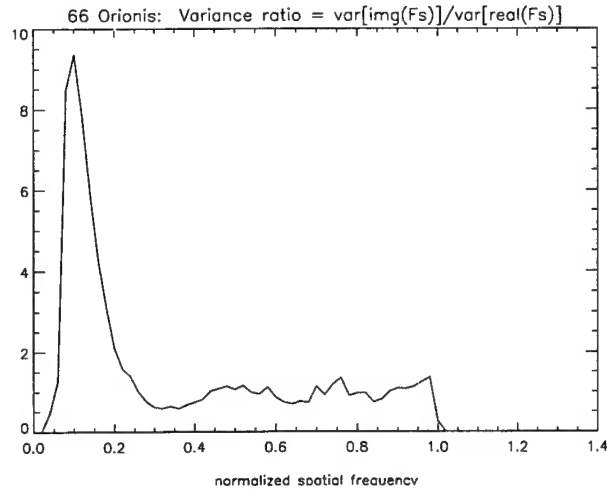
Figure 5 Variance ratio of the 66 Orionis image
Ratio of the variance of the imaginary component of the
Fourier spectrum of the 66 Orionis image to the variance of
the real component.

SOR MAMA data: The data sets taken at the SOR have varied spectral bandpasses as
needed for the power spectrum SNR theory validation. To be useful for this purpose, however,
the power spectrum SNR must meet the quality requirements; namely, it must reveal a significant
speckle imaging wing. For the star SAO 68166, Figure 6 and Figure 7 show the radial average of
the power spectrum and its SNR over 1784 short exposures of 9.6ms time duration. It is known
that D=1.5m and λ=550nm, but the frequency axis is not normalized because the diffraction limit
is unknown due to the absence of scaling information for this group of data. It should also be
mentioned that the plots zoom in to frequencies 0-100 although the data actually ranges to 256.
In analyzing this power spectrum, it is evident that the speckle imaging wing is nearly
nonexistent (begins at spatial frequency 5 and rapidly tapers off). Another data set, obtained with
the same camera and system, was processed for reference star 6995. Unfortunately, these results
(not shown) were even worse than with SAO 68166 in that the speckle imaging wing at mid
frequencies is almost non-existent.

CCD data: The third group of short exposure, photon-limited, unresolved star data was
taken with a CCD camera at AMOS in March of 1992. We analyzed images for the reference
stars ASN 5747 (magnitude 3.67) and ASN 5953, but will only present results of the former as

32-11
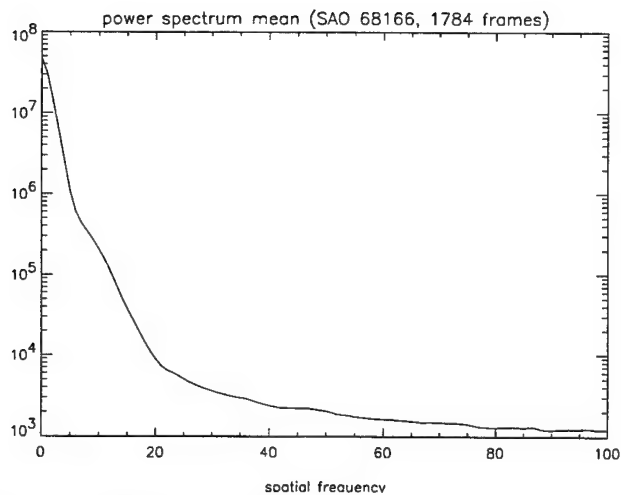
Figure 6 Radial average power spectrum of SAO 68166
Radial average of the mean power spectrum of the
reference star SAO 68166 taken at the SOR, June 1993.
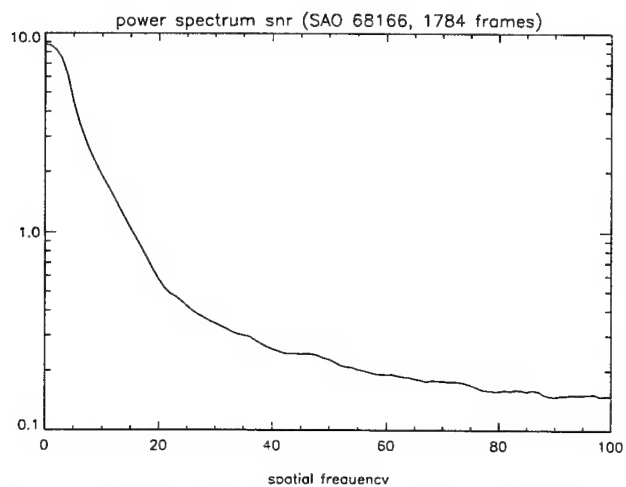
Figure 7 Power spectrum SNR of SAO 68166
SNR of the radial average power spectrum of the reference
star SAO 68166 taken at the SOR, June 1993.

representative of both. This data set is fixed at 30ms exposures (slightly longer than a true "short" exposure) and there are only approximately 350 frames per data set. The spectral bandpass is fixed at 140nm, D=1.6m, and λ is approximately 700nm. As with the Hege data, we analyzed this data group to calculate the Fourier spectrum noise characteristics. First, we processed the flat fields and dark frames that accompanied the data. Then we adjusted the raw image data with the average flat and dark fields in order to mitigate the multiplicative and additive artifacts of the camera's recording process.

In calculating the power spectra and bispectra, we assumed that the read noise was simply the standard deviation of the image pixel values of the average dark frame. With this assumption, however, the power spectrum of the star ASN 5747 reveals two large bands indicating that noise is still present (Figure 8). Changing the assumed read noise to be the average power spectrum of the dark frame appears to provide the correct, unbiased power spectrum (Figure 9). This adjusts the image at each pixel individually instead of the assumption that a constant subtraction suffices. With this change, the mean power spectrum, power spectrum SNR, estimated phase and Fourier spectrum variance ratio for the reference star ASN 5747 are shown in Figure 10 through Figure 13. The power spectrum (Figure 10) demonstrates a very slight mid frequency speckle imaging wing unlike the pronounced wing of the Hege data. This is probably due to attenuation from astigmatism which would equally attenuate both the real and imaginary components of the Fourier spectrum, so this phenomenon does not rule out the usefulness of this data group. The power spectrum SNR (Figure 11) shows reasonably high values up until approximately 0.3. The SNR at mid frequencies might be higher if more frames were available to be included in the processing.
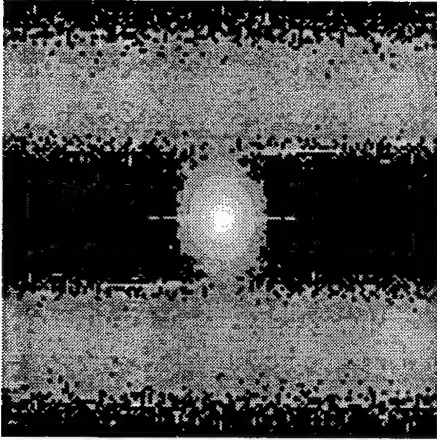
Figure 8 The average power spectrum of ASN 5747 calculated with the assumption of constant read noise
A picture of the average power spectrum of the reference star ASN 5747 taken at the Air Force Maui Optical Site (AMOS) in March 1992. This picture illustrates the power spectrum as calculated under the incorrect assumption that the read noise of this CCD camera is a constant among all pixel locations.
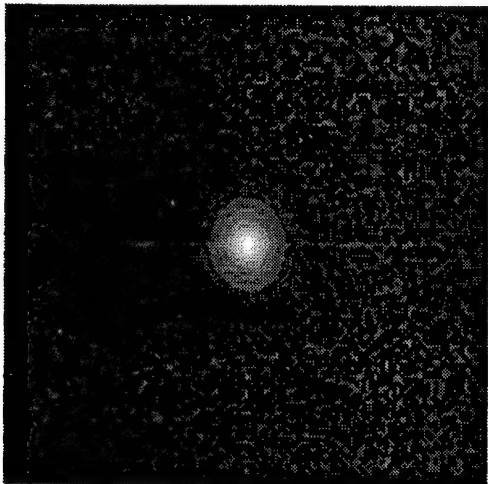
Figure 9 The average power spectrum of ASN 5747 calculated with the estimated per pixel read noise
A picture of the average power spectrum of the reference star ASN 5747 taken at the AMOS in March 1992. This picture illustrates the power spectrum as calculated by accounting for the read noise of each pixel individually.

radial average of power spectrum (Ref. star ASN 5747, 353 frames)

normalized spatial frequency
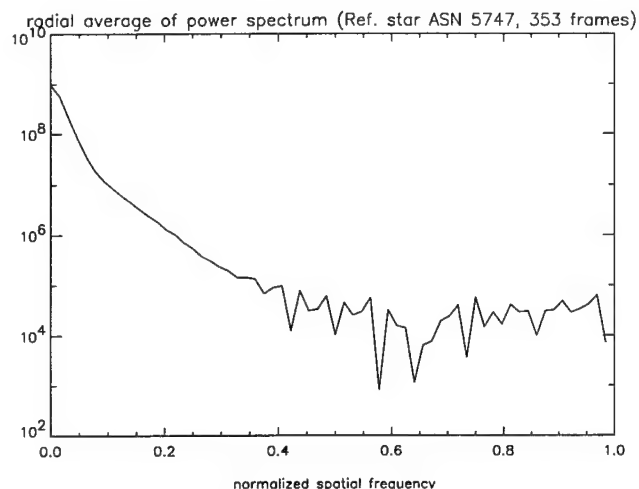
Figure 10 The radial average of the power spectrum of ASN 5747 calculated with estimated per pixel read noise
The radial average of the power spectrum of the reference star ASN 5747 taken at the AMOS in March 1992. This shows the power spectrum as calculated by accounting for the read noise of each pixel individually.

radial average of power spectrum snr (Ref. star ASN 5747, 353 frames)

normalized spatial frequency
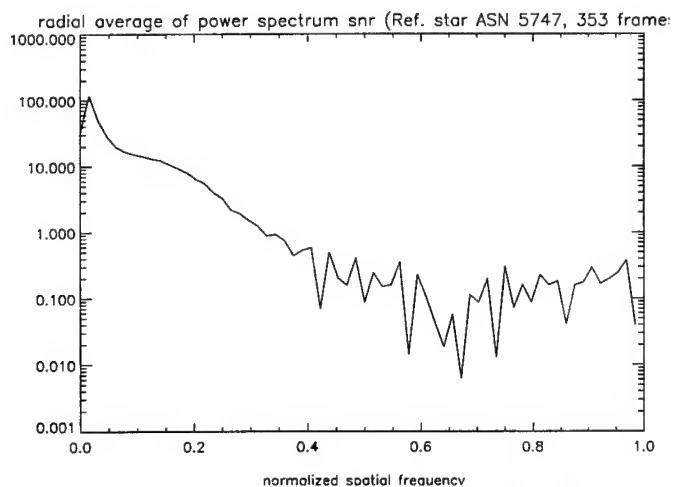
Figure 11 Power spectrum SNR of ASN 5747
SNR of the radial average power spectrum of the reference star ASN 5747 taken at the AMOS in March 1992 using a power spectrum as calculated by accounting for the read noise of each pixel individually.
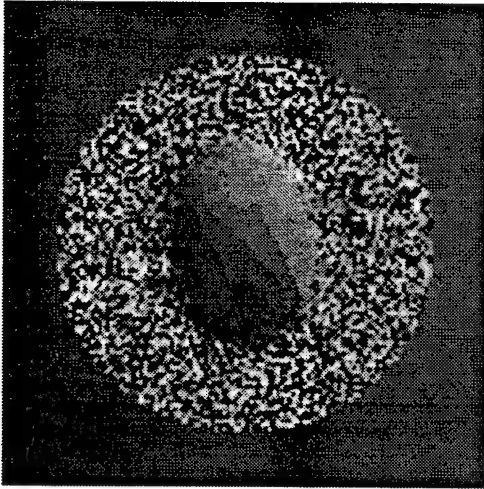
Figure 12 Estimated Phase of the ASN 5747 image
Picture of the estimated phase of the reference star ASN
5747 as calculated with the power spectrum adjusted for
per pixel read noise, but a constant read noise in the
bispectrum. Black corresponds to -π and white to +π.

This group does have an anomaly, though: Figure 11 shows an increasing slope at d.c. whose
cause is unknown. Also, as seen in Figure 12, the phase plots are ellipsoidal rather than spherical,
indicating a possible dispersion correction problem. The point should be made, however, that the
phase and Fourier spectrum are obtained via the bispectrum whose calculation has not been
adjusted for the non-ideal read noise. Therefore, any conclusions based on the estimated phase
or Fourier spectrum are premature. At this time we have not determined the necessary changes
that need to be made to the bispectrum calculations. Despite the incomplete read noise
adjustment, the Fourier spectrum variance ratio (Figure 13) appears to be a potential candidate
for prior knowledge application since the ratio meets the required separation factor of 0.6 up until
spatial frequencies of approximately 0.25. Therefore, we anticipate that when the noise
correction has been made in the bispectrum, this data will be useful for application and
investigation of Matson's prior knowledge theories.
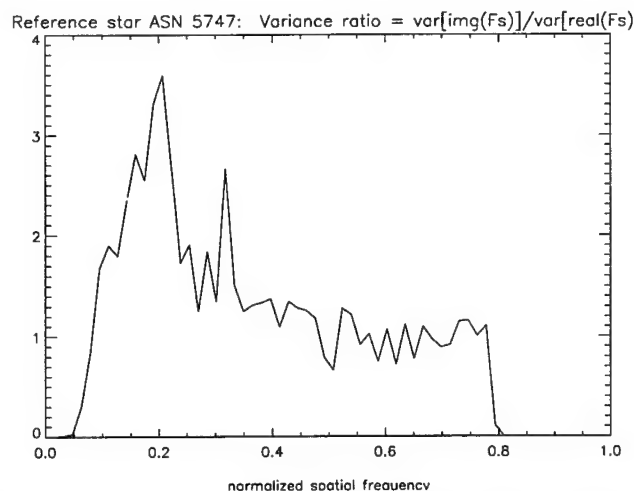
Figure 13 Variance ratio of the ASN 5747 image
Ratio of the variance of the imaginary component of the
Fourier spectrum of the ASN 5747 image to the variance of
its real component. The Fourier spectrum was calculated
using a per pixel read noise adjustment in the power
spectrum, but a constant read noise in the bispectrum.

4. Conclusions

The lack of a speckle imaging wing in the power spectra of the SOR MAMA data indicates
that the camera may have been defocused. This phenomenon underlines the need to calculate
and view the power spectra on-site at the time of data acquisition to ensure the data will be
applicable to the purpose for which it is intended. Unfortunately, the defocused nature of the
SOR MAMA data excludes it from being useful for our intent to investigate the effect of spectral
bandwidth on the image's power spectrum SNR. Since this data set is the only one employing
varying spectral bandwidths, we do not presently have other data with which to explore the
theory. We anticipate procurement of appropriate data in the next year, and hope to perform the
intended research next summer.

The AMOS CCD data will likely be useful for exploring the prior knowledge theories after
the appropriate read noise corrections have been made to the bispectrum calculations. The Hege
data, however, presently meets the requirements for examining these theories. Our future plans
include looking into these results more carefully.

REFERENCES

1. C. L. Matson, Variance reduction in Fourier spectra and their corresponding images with the use of support constraints. *J. Opt. Soc. Am. A* **11**, 97-106 (1994).
2. D. L. Fried and D. L. Hench, Spectral bandwidth and adaptive optics considerations for white light speckle imagery. *Optical Sci. Co.*, rep. TR-1033, 1990.
3. A. Labeyrie, Attainment of diffraction-limited resolution in large telescopes by Fourier analyzing speckle patterns in star images. *Astr. Astrophys.* **6,** 85-87 (1970).
4. A. W. Lohmann, G. P. Weigelt and B. Wirnitzer, Speckle masking in astronomy: triple correlation theory and applications. *Appl. Opt.* **22,** 4028-4037 (1983).
5. C. L. Matson, Weighted-least-squares phase reconstruction from the bispectrum. *J. Opt. Soc. Am. A* **8**, 1905-1913 (1991).

TOPAZ II REACTOR CONTROL LAW IMPROVEMENT

Yuri B. Shtessel
Associate Professor
Department of Electrical and Computer Engineering

University of Alabama in Huntsville
Huntsville, Alabama 35899

TOPAZ II REACTOR CONTROL LAW IMPROVEMENT

Yuri B. Shtessel
Associate Professor
Department of Electrical and Computer Engineering
University of Alabama in Huntsville
Huntsville, Alabama 35899

Abstract

The 6 kW_e space nuclear reactor power system TOPAZ II that generates electricity from nuclear heat using in-core thermionic converters is considered. The major subsystems that comprise the power system are: the nuclear reactor, which contains the thermionic converters, the radiation shield, the reactor coolant system, the cesium supply system, the instrumentation and automatic control system. The TOPAZ II Automatic Control System is mainly designed to start up the TOPAZ II system, to control the reactor thermal and electric power being supplied to the spacecraft payload and to shut down the TOPAZ II system.

The control algorithms of the TOPAZ II system has been improved to meet the following main goal. This is *to ensure the reduced sensitivity of the reactor thermal and electric power control processes to the system parameter uncertainties mostly associated with the neutron detectors ("ionization chambers") and external disturbances while providing the reference electric or thermal power profile following.*

The enhanced controllers have been designed via implementation of the sliding mode control technique which is very helpful in the controller design upon the uncertain environment. The TOPAZ II Automatic Control System with the designed sliding mode controllers has been simulated in the start up mode and the electric power (current) profile following mode. The result of simulations showed the reduced sensitivity of the control processes to the variation of the parameters of the system and the improved accuracy and robustness of the thermal and electric power (current) reference profile following. The realization of the designed sliding mode controllers in the TOPAZ II Automatic Control System

- should improve the accuracy of the system performance,
- should bring more reliability and flexibility to the TOPAZ II operation.

TOPAZ II REACTOR CONTROL LAW IMPROVEMENT

Yuri B. Shtessel

1. Introduction

The TOPAZ II is a 6 kW_e nuclear reactor power system that generates electricity from nuclear heat using in-core thermionic converters [1,2]. The TOPAZ II has been designed in Russia for operations in space and should help to establish the infrastructure required for the launch and use of future US space reactor systems. The major subsystems that comprise the power system are [1,2]: a) the nuclear reactor, which contains the thermionic converters; b) the radiation shield; c) the reactor coolant system; d) the cesium supply system; e) the instrumentation and control (I&C) system.

The nuclear reactor contains 37 single-cell thermionic fuel elements (TFE), which are fueled by uranium dioxide (UO_2) fuel pellets that are 96% enriched in U^{235}. Three of TFEs are used to drive the electromagnetic (EM) pump and the remaining 34 provide power to operate the TOPAZ II reactor and the satellite payload. The TFEs are set within channels in block of $ZrH_{1.85}$ moderator, which is canned in stainless steel. The reactor core is surrounded by radial and axial beryllium (Be) reflector. The radial reflector contains three safety drums and nine control drums. Each drum contains a section of boron carbide (B_4C) neutron poison on its periphery that is used to control the nuclear reaction by drum rotation. The safety drums operate at fixed rotational speed of 22.5^o/sec, whereas the nine control drums operate at a variable rotational speed of up to 1.4^o/sec. During the operation, the nuclear fuel heats the TFE emitters to 1800° to 2100° K. The waste heat is removed by the coolant system, which flows past the outer surface of the collector of the thermionic unit and maintains the collector at ~900^oK during operation.

The radiation shield is attached to the lower end of the reactor. The shield is composed of a stainless steel shell that contains lithium hydride (LiH).

The reactor coolant system includes sodium-potassium (NaK) coolant, a single EM pump, stainless steel piping and a heat rejection radiator.

The cesium supply system provides cesium (Cs) to the TFE inter electrode gap. Cesium is necessary to suppress the space charge that occurs near the emitters of thermionic converters; suppressing this charge increases the efficiency of the converter.

The TOPAZ II instrumentation and control system (I&C) provides the means of monitoring, controlling and displaying system conditions. The control components employed to manage these activities are grouped under the name of Automatic Control System (ACS). The ACS is a closed-loop feedback control system. It controls reactor system operations by receiving signals from various sensors (thermocouples, flux detectors, current shunts, ionization chambers, etc.) located on the reactor system, uses this information to calculate the required actions necessary to put or maintain system parameters within their specified, and sends out command signals to perform the computed adjustments.

The principles that guided the ACS design effort centered on the need for an automatic control system capable of performing the various functions. The main functions are follows:

- Start up the TOPAZ II system.
- Maintain the operations of TOPAZ II in nominal conditions.
- Stabilize the voltage supplied to the payload.
- Provide the following of a payload reference electric power (current) trajectory.
- Maintain safety control.
- Shut down the TOPAZ II system.

The ACS is mainly designed to control the reactor thermal and electric power being supplied to the spacecraft payload within the following ranges (and accuracy):

- reactor thermal power: 0-150% ($\pm$ 700 watts),
- electric power supply: 30-125% ($\pm$ 3.5 amperes, $\pm$ 0.8 volts).

The ACS has three functional systems, namely the Telemetry and Command System (TCS), the Automatic Regulating System (ARS) and the (Electric) Power Supply System (PSS). The fig 1. [1,2,3] gives a generalized simplified overview of the functional diagram of the TOPAZ II nuclear reactor system.
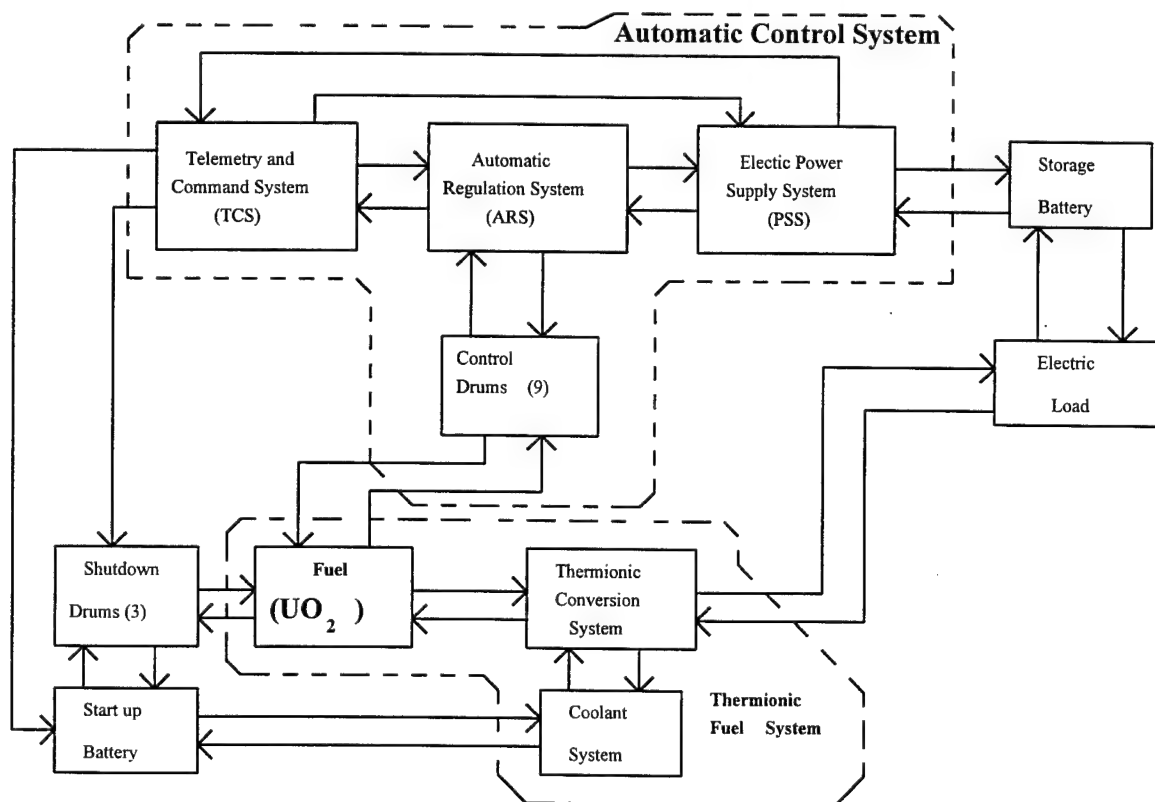
Fig. 1. Functional diagram of the TOPAZ II nuclear reactor system

In this work we should *develop the TOPAZ II reactor control system to ensure the reduced sensitivity of the reactor control processes to the system parameter uncertainties mostly associated with the neutron detectors ("ionization chambers") and external disturbances while providing the reference electric or thermal power profile following.*

During the development of the reactor ARS we should solve the following problems:

♥   *The first problem* is to provide the robust control of the start-up mode of the TOPAZ II reactor system via sliding mode controller.

♥   *The second problem* is to provide the robust reference payload electric current (power) profile following via corresponding sliding mode controller. This operation mode of the TOPAZ II should be a new one.

These problems have been solved in this work via implementation of the sliding mode control technique which is very helpful in the controller design upon the uncertain environment. The realization of the designed sliding mode controllers in the TOPAZ II ARS should

- improve the accuracy an robustness of the system performance,
- bring more reliability and flexibility to the TOPAZ II operation and
- reduce the dissipation of the heat through the shunt current/voltage controller in the PSS.

2. Methodology of the Sliding Mode Control

Let us consider a single-input-single-output linear control system. This is

$$\begin{cases} \dot{x} = Ax + bu + f(t) \\ y = gx, \end{cases} \tag{1}$$

where $x \in \mathrm{R}^n$ - a state vector, $f \in \mathrm{R}^n$ - an external vector-disturbance, $u \in \mathrm{R}^1$ - a control function, $y \in \mathrm{R}^1$ -an output, $A \in \mathrm{R}^{n \times n}$ - a constant matrix, $b \in \mathrm{R}^n$ - a constant column-vector, $g \in \mathrm{R}^n$ - a constant row-vector, $\{A, b\}$ - a completely controllable pair.

The objective is to design the control function u to provide $\lim y(t) = 0$ as $t \to \infty$ and to make this transient process desirable.

The solution of this control problem in sliding mode implies [4,5,6]

- the synthesis of the sliding surface $\sigma = 0, \sigma \in \mathrm{R}^1$ in the in the subspace of output and its derivatives (or in some state-variable space) such the behavior of the output of the system (1) in this surface will be desired,
- to force a control system to move to the sliding surface and maintain this system in the sliding surface thereafter by high speed switching control functions.

2.1 The Sliding Surface Synthesis in the Subspace of the Output and Its Derivatives

Differentiating the output of the system (1) k times until $gA^{k-1}b \neq 0$ we will transform this system to the following canonical form [4,5]

$$\begin{cases} \dot{y}_1 = y_2, \quad \dot{y}_2 = y_3, \dots, \dot{y}_k = gA^k x + \sum_{j=0}^{k-1} A^j f^{(k-j-1)}(t) + gA^{k-1} bu, \\ y = y_1 \end{cases} \tag{2}$$

The following surface in the k-dimensional subspace can be formed

$$\sigma = \sum_{i=1}^{k} c_i y_i = 0, \quad c_k = 1. \tag{3}$$

To find the equations of the sliding mode in the system (2) in the surface (3) we need [4,5] to solve the equation (3) for y_k and to substitute obtained solutions in equations (2), then to skip the last differential equation in the system (2). That is

$$\begin{cases} \dot{y}_1 = y_2, \quad \dot{y}_2 = y_3, \ldots, \dot{y}_{k-1} = -\displaystyle\sum_{j=0}^{k-1} c_j y_j, \\ y = y_1 \end{cases} \tag{4}$$

The system (4) can be rewritten as one differential equation of the k-th order as follows

$$y^{(k-1)} + c_{k-1} y^{(k-2)} + \cdots + c_2 \dot{y} + c_1 y = 0 \tag{5}$$

The values of coefficients $c_1, c_2, \ldots, c_{k-1}$ could be assigned to provide the desired eigenvalues placement of the characteristic equation of the differential equation (5) [4,5].

Remark 1. The motion of the output y in the sliding mode depends only on coefficients $c_1, c_2, \ldots, c_{k-1}$ of the sliding surface (3) and is invariant to the parameters A, b, g of the plant (1) and to the external disturbance $f(t)$.

Remark 2. One needs to differentiate the output of the system k times to realize the sliding mode controller. This differentiation should make a big problem upon the noisy measurements of the output y. This multiple differentiation can be avoided by means of the calculation of the output's derivatives through the state variables of the system (1). In this case the motion of the system (1) in sliding mode (5) should depend on some parameters of the system (1).

2.2 Discontinuous Control Function Design

Let us consider the following discontinuous control law

$$u = \begin{cases} u^+, \text{if } \sigma > 0 \\ u^-, \text{if } \sigma < 0 \end{cases} \tag{6}$$

The functions u^+, u^- should be specified to make the sliding surfaces (3) or (7) "attractive" for the trajectories of the system (1). This condition [4,5] $\sigma \cdot \dot{\sigma} < 0$ implies the corresponding inequalities which have to be met by means of the appropriate specification of the functions u^+, u^-.

Remark 3. Based on the above statements we can conclude that in Sliding Mode the system response is insensitive to certain parameter variations, nonlinearities and external and interconnection disturbances [4,5] which makes the sliding mode control strategy attractive for the output tracking problem solution in nonlinear systems with uncertainties and disturbances.

3. The mathematical model of the TOPAZ II

The mathematical model of the TOPAZ II system consists of several sub models. They are [3,7].

3.1 Neutronic model

The Topaz II power system employs a small and compact core, in which neutronic responses of the reactor are tightly coupled. The changes occur in the reactor may be described by treating the entire core as a "point". The neutronic model describing the response of the reactor can be taken as follows [3,7]:

$$\begin{cases} \dfrac{dQ}{dt} = \dfrac{\rho(t) - \overline{\beta}}{l^*} Q + \sum_{i=1}^{6} \lambda_i C_i + Q^*(t) \\ \dfrac{dC_i}{dt} = \dfrac{\beta_i}{l^*} Q - \lambda_i C_i \quad \forall i = 1,6 \end{cases} \tag{7}$$

where

Q is the weighted integral of all neutrons present in the reactor normalized to the heat power,

$\rho(t)$ is the net reactivity of the reactor, $Q^*(t)$ is the effective neutron source strength,

l^* is the prompt neutron lifetime, λ_i is the decay constant for the i-th precursor group,

C_i is the concentration of the i-th precursor group normalized to the initial power,

β_i is the fractional yield of the i-th group of delayed neutrons.

     The respective kinetics parameters are given as follows

$$\begin{cases} \beta_1 = 3.040 \cdot 10^{-4}, \quad \beta_2 = 1.704 \cdot 10^{-3}, \quad \beta_3 = 1.504 \cdot 10^{-3}, \\ \beta_4 = 3.256 \cdot 10^{-3}, \quad \beta_5 = 1.024 \cdot 10^{-3}, \quad \beta_6 = 2.080 \cdot 10^{-4}, \\ \lambda_1 = 0.0127\, s^{-1}, \quad \lambda_2 = 0.0317\, s^{-1}, \quad \lambda_3 = 0.1160\, s^{-1}, \\ \lambda_4 = 0.3110\, s^{-1}, \quad \lambda_5 = 1.3970\, s^{-1}, \quad \lambda_6 = 3.8720\, s^{-1}, \\ \overline{\beta} = 0.008, \quad\quad\quad l^* = 2 \cdot 10^{-5}, \quad\quad Q^*(t) = 1.46 \cdot 10^{-5}\, W/s. \end{cases} \tag{8}$$

The initial values of the coordinates of the system (7) have been selected as follows:

$Q(0) = 100000\,W$ or $115000\,W$, $C_1(0) = 1.197\ 10^8$, $C_2(0) = 2.688\ 10^8$, $C_3(0) = 6.483\ 10^7$,

$C_4(0) = 5.235\ 10^7$, $C_5(0) = 3.665\ 10^6$, $C_6(0) = 2.686\ 10^5$.   (9)

The net reactivity of the nuclear reactor is given in the form:

$$\rho(t) = \rho_c(t) + \Delta\rho(t), \tag{10}$$

where $\rho_c(t)$ is the integral reactivity worth of the control drums, and $\Delta\rho(t)$ is the reactivity feedback of the reactor. The integral reactivity worth of the control drums is given as follows [3,7]:

$$\rho_c(t) = 6.89 \cdot 10^{-13} \Theta^5 - 2.33 \cdot 10^{-10} \Theta^4 + 3.28 \cdot 10^{-9} \Theta^3 + 4.57 \cdot 10^{-6} \Theta^2 - 5.88 \cdot 10^{-5} \Theta + \beta, \tag{11}$$

where Θ is the angular position of the main control drum expressed in degrees.

Two values of β have been used for the reactor kinetic description:

- $\beta = -2.3 \cdot 10^{-2}$, which implies the controlled reactivity range $-2.875 \le \rho_c(t) \le 2.398$ in \$,
- $\beta = -3.5 \cdot 10^{-2}$, which implies the controlled reactivity range $-4.375 \le \rho_c(t) \le 0.9875$ in \$.

The reactivity feedback of the reactor caused by temperature change in the ZrH moderator, reflector, reactor core support plates, UO_2 fuel and TFE electrodes have been determined under isothermal conditions up to a temperature of 1000 K. The main component of the reactivity feedback of the reactor is determined by the moderator. This is:

$$\begin{cases} \dfrac{dT_m}{dt} = -4 \cdot 10^{-3} T_m + 2 \cdot 10^{-5} Q + 1.2, \\ \Delta \rho = -8.22 \cdot 10^{-16} (T_m - 300)^5 + 1.6 \cdot 10^{-12} (T_m - 300)^4 - 1.11 \cdot 10^{-9} (T_m - 300)^3 + \\ \qquad + 2.92 \cdot 10^{-7} (T_m - 300)^2 + 1.76 \cdot 10^{-5} (T_m - 300) \end{cases} \tag{12}$$

The initial value of T_m has been selected as follows: $T_m = 800\,K$ for the current profile following mode and $T_m = 300\,K$ for the start up mode.

3.2 Thermal Model

Among the thermal processes in the reflector, reactor core support plates, the UO_2 fuel and TFE electrodes the thermal process in the UO_2 fuel is the "fastest" one. This process in the UO_2 fuel is described as follows [3,7]:

$$\frac{dT_f}{dt} = -0.0125 T_f + 2.295 \cdot 10^{-4} Q + 3.75. \tag{13}$$

The initial value of $T_f = 2136\,K$ for the current profile following mode and $T_f = 300\,K$ for the start up mode.

3.3 Thermionics Model

We assume that the output voltage is stabilized by a separate controller at 27 volts level and the cesium pressure is stabilized at the level 2 torr. Then the payload current I is described by the following differential equation [3,7]:

$$\frac{dI}{dt} = -4.07 \cdot 10^{-4} I + 1.874 \cdot 10^{-5} Q. \tag{14}$$

The description of the output electric power P can be also very useful. This is

$$\frac{dP}{dt} = -0.011 P + 5.56 \cdot 10^{-4} Q. \tag{15}$$

Two sets of the initial values been selected as follows:

$$I(0) = 184.999\,A, \quad P(0) = 5000\,W \text{ and } I(0) = 213.8\,A, \quad P(0) = 5750\,W. \tag{16}$$

3.4 Control Drum Drive Model

The control drum drive system [1,2] is designed to rotate the reactor control drums from 0° to the angle < 180°. Angular velocity of the drums is 1.4°/sec when the motor is operating at 135 Hz. The stepper motor speed is 500 rpm at 135 Hz and is 0.1°/sec when the motor is operating at 9 Hz. We will use the equivalent mathematical model of a stepper motor as an ordinary differential equation of a second order. This is [8]

$$\begin{cases} \dfrac{\pi}{180} (J_m + n^2 J_L) \ddot{\Theta}_m = T_m - n T_L, \\ T_m = \dfrac{k_i}{R_a} U, \\ \Theta = n \Theta_m, \quad \Omega = n \Omega_m, \quad \dot{\Theta}_m = \Omega_m. \end{cases} \tag{17}$$

where

$J_m, J_L \ (Nm \sec^2/rad)$	are the inertia moments of the motor and load (drums) correspondingly,
$T_m, T_L \ (Nm)$	are the torques of the motor and a load (disturbance) correspondingly,
n	is a gear ratio $(n=1{:}2102)$,
$\Theta_m, \Theta \ (degrees)$	are the position angles of the motor and the drum correspondingly,
$\Omega_m, \Omega \ (degrees/\sec)$	are the angular velocities of the motor and the drums correspondingly,
$U \ (volts)$	is the equivalent control voltage,
$k_i \ (Nm/A), \ R_a \ (Om)$	are the torque constant and the armature resistance of the motor correspondingly.

The control drum drive parameters have been taken as follows

$$\begin{cases} k_i = 0.0741 \, Nm/A, & R_a = 10 \, Om, \\ J_m = 2\cdot10^{-3} \, Nm\sec^2/rad, & J_L = 7.4\cdot10^3 \, Nm\sec^2/rad. \end{cases} \tag{18}$$

The resulting equivalent description of the control drum drive has been obtained as follows

$$\begin{cases} \dot{\Theta} = \Omega, \\ \dot{\Omega} = 0.055U - 3.52\cdot10^{-3} T_L, \end{cases} \tag{19}$$

while $|\Omega| \le 1.4^0/\sec, 0 \le \Theta < 180^0$.

The following initial values of parameters and the disturbance torque have been selected for the simulation purpose:

$$\Theta(0) = 90 \, degrees, \quad \Omega(0) = 0,$$

$$T_L = 85.12[1(t-75) - 1(t-95)] - 170.24[1(t-95) - 1(t-120)] \, Nm. \tag{20}$$

3.5 Payload Reference Electric Model

We assume the reactor start-up mode is finished. It means that the reactor is to be in a nominal operational mode already. The desired change (profile) of the reference payload current from the nominal value can be caused by the scheduled turn on or turn off some payload unit. This change has been defined as follows [3,7]:

$$\tau \frac{dI_{ref}}{dt} = -I_{ref} + I_{ref}^{fin}, \tag{21}$$

where τ is a time constant; I_{ref}^{fin} is a final value of the reference payload current change.

The desired change (profile) of the reference payload electric power from the nominal value can be specified similarly. This is [3,7]

$$\tau \frac{dP_{ref}}{dt} = -P_{ref} + P_{ref}^{fin}, \tag{22}$$

where τ is a time constant; P_{ref}^{fin} is a final value of the payload electric power change.

We will specify the time constant τ upon the assumption that the immediate or average rate of change with respect to time of the neutronic (heat) power of the reactor is limited. This is

$$\left| \frac{dQ}{dt} \right| \le 80 \, W/\sec. \tag{23}$$

The inequality (23) should be met for any moment at the control interval if the time constant satisfies the following inequality

$$\tau \geq 6.75(I_{fin} - I_{in}).$$ (24)

If to think about the inequality (23) to be met in average at the control interval than

$$\tau \geq 1.68(I_{fin} - I_{in}),$$ (25)

where I_{in}, I_{fin} is the initial (nominal) and final current values correspondingly.

The following values of parameters have been specified during the simulations:

$$I_{in} = 185\,A \ (or \ 213.8\,A), \quad I_{fin} = 206\,A \ (or \ 234.8\,A), \quad \tau = 30\,\text{sec}$$ (26)

3.6 Start Up Thermal Power Reference Profile

The TOPAZ II reactor start up mode consists of the following operations [1,2]

- Rotate shutdown drums to 180° (full out).
- Rotate control drums at full speed (500 revolutions/min of the stepper motor $\Rightarrow$ 1.4°/sec of control drums).
- When the position angle of control drums is $\approx 154°$ then reverse motor direction until the position angle is $\approx 145°$ then stop.
- When the reactor thermal power is ≈ 6 kW then begin feedback control to maintain a 600 W/s power ramp rate.
- When the reactor thermal power is ≈ 30 kW, adjust a reactor power ramp rate to 80 W/s until reactor power is 115 kW, then maintain 115 kW while system is achieving steady state mode.

4. Control Problem Formulations

Two control problems have been formulated and solved in this work. They are.

4.1 The Reference Thermal Power Profile Following (Start Up Mode)

The following assumptions have been done

- the TOPAZ II reactor system is considered to be described by the equations (7) - (26),
- first three steps of the start up mode (without feedback control) is over and the reactor thermal power is ≈ 6 kW.

The problem is to specify the parameters U^+, U^- and σ of the discontinuous control law

$$U = \begin{cases} U^+, & \text{if } \sigma > 0, \\ U^-, & \text{if } \sigma < 0, \end{cases}$$ (27)

where U^+, U^- are the continuos functions of the coordinates of the reactor system and σ is a sliding surface, to provide the desired following of a reference thermal power profile in the sliding surface $\sigma = 0$.

4.2 The Reference Current Profile Following

The following assumptions have been done:

- the TOPAZ II reactor system is considered to be described by the equations (7) - (26),
- the start up mode has been over.

The problem is to specify the parameters U^+, U^- and σ of the discontinuous control law (27) to provide the desired following of a payload reference electric power (current) profile in the sliding surface $\sigma = 0$.

5. The Design of TOPAZ II Reactor Sliding Mode Controllers.

At the first step we should specify the equation of the sliding surface $\sigma = 0$ so that the thermal power tracking error $e = Q - Q_{ref}$ (payload current tracking error $\varepsilon = I - I_{ref}$ or electric power tracking error $\varphi = P - P_{ref}$) has the desired performance in the sliding surface, where Q_{ref} is a reference thermal power profile, I_{ref} (P_{ref}) is a payload reference current (electric power) profile. For example, we can think about the tracking error performance as a desired one if the tracking error is described by the ordinary linear time-invariant homogeneous differential equation with the desired (given) eigenvalues placement.

At the second step we plan to specify discontinuous control functions U^+, U^-

- to provide the movement of the reactor system to this sliding surface,
- to maintain the system in the sliding surface thereafter.

5.1 Methodology of the TOPAZ II Reactor Sliding Mode Controller Design

First, we should look for a sliding surface equation as a linear form of the tracking error and its derivatives. This is

$$\sigma = \sum_{k=0}^{N} \alpha_k e^{(k)} = 0, \quad \alpha_N = 1. \tag{28}$$

We need to define the value N and the values α_k $\forall k = 0, N-1$ to specify the equation of a sliding surface (28).

To specify the value N we need to

- differentiate the tracking error K times so that its K-*th* derivative contains the control function U,
- specify N as follows: $N = K - 1$.

We should specify the values α_k $\forall k = 0, N-1$ to make the solution to the differential equation (28) to be the best in some meaning. We will think about the solution to the differential equation as the best one if it minimizes the *integral of time multiplied by absolute error* (ITAE) [9]:

$$J = \int_0^{T_S} t|e(t)|dt, \tag{29}$$

where T_S is the settling time which is the time required for the system to settle within a 2% of an input amplitude.

Second, we should look for the continuous control functions U^+, U^- which provide the movement of the reactor system to the sliding surface (28) and maintain the system in the sliding surface thereafter. The conditions for the specification of the control functions U^+, U^- can be derived on the base of the well known inequality [4,5]

$$\sigma \cdot \dot{\sigma} < 0, \tag{30}$$

which expresses the existence condition of the sliding mode in the sliding surface.

5.2 The Sliding Mode Controller Design for the Reactor Start Up Mode

First, differentiating the tracking error $e = Q - Q_{ref}$ through the system (7) - (26) we can find that the third derivative $e^{(3)} \sim U$ and $N = 2$ in the equation of the sliding surface (28). It means that the sliding surface (28) actually is the differential equation of the second order. The characteristic equation of the differential equation of the second order (28) can be formed as follows:

$$r^2 + 2\xi\omega_n r + \omega_n^2 = 0, \tag{31}$$

where ω_n is to be defined upon requirements to the settling time of the transient process in the sliding mode. This is

$$\omega_n \geq 4/\xi T_s. \tag{32}$$

The optimal value of ξ in expression (32) is known [9] to be equal to 0.7.

Second, the optimal coefficients of the differential equation (28) with $N = 2$ should be specified as follows

$$\alpha_o = \omega_n^2, \quad \alpha_1 = 2\xi\omega_n. \tag{33}$$

The following value of the settling time has been selected: $T_s = 10\,\text{sec}$ and the corresponding value of ω_n has been specified as follows: $\omega_n = 0.57$. The coefficients (33) became:

$$\alpha_0 = 0.327, \quad \alpha_1 = 0.8. \tag{34}$$

Third, the following values of U^+, U^- in discontinuous control law (27) were specified:

$$U^+ = 27 \text{ volts}, \quad U^+ = \text{-27 volts} \tag{35}$$

to meet the inequality (30) calculated through the equations of the system (7)-(26).

The block-diagram of the sliding mode controller (27), (28), (34), (35) is depicted at the fig. 2.
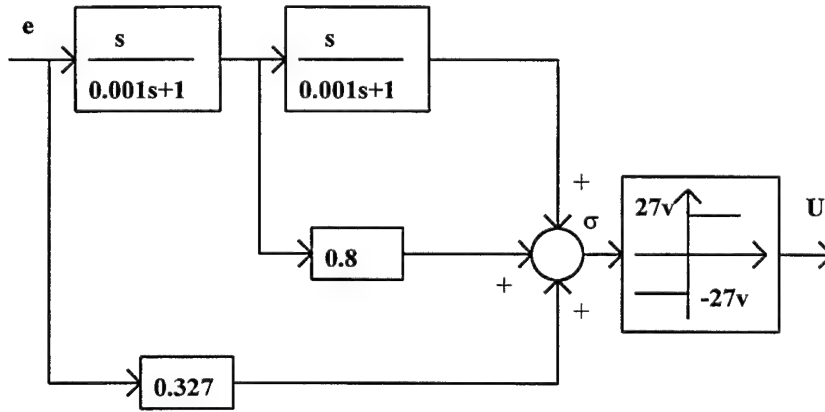
Fig. 2 The block-diagram of the start up sliding mode controller

Remark 4. The sliding mode controller shown at the fig. 2 is related to the actuator in the form of DC motor, described by the equations (17)-(20). Actually the stepper motor is the actuator in the TOPAZ II reactor system. The designed sliding mode controller needs adjustment to be implemented in the system with a stepper motor. It is obvious from the equations (19) that

$$\Omega = \int_0^t [0.055U - 0.00352T_L)d\tau, \quad \Omega_m = \frac{1}{n}\Omega, \tag{36}$$

$|\Omega| \leq 1.4$ %/sec and $|\Omega_m| \leq 3000$ %/sec, then the signal (36) has to be transformed by a multiplexer to the pulses with 27 V magnitude and the frequency proportional to the signal (36) so that 1.4 %/sec $\Rightarrow 138$ Hz [1,2,3]. The block-diagram of the adjusted start up sliding mode controller upon the assumption $T_L = 0$ is depicted at the fig. 3.
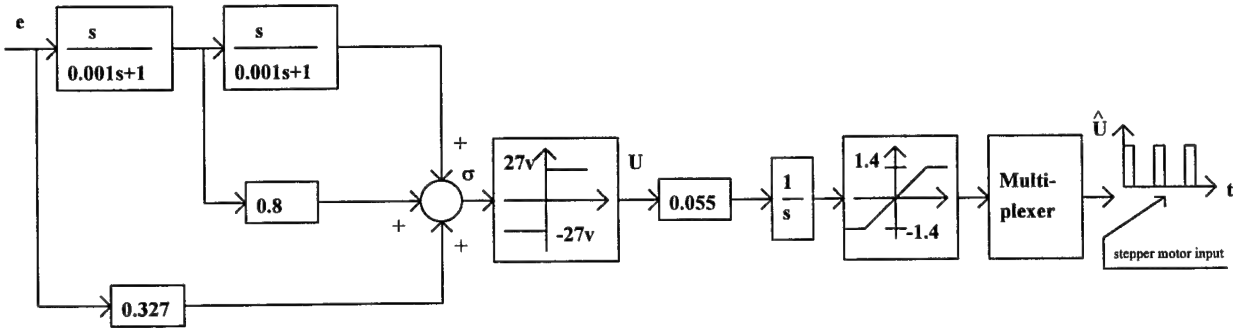
Fig. 3 The block-diagram of the adjusted start up sliding mode controller

Let us study the sensitivity of the start up sliding mode controller to the gain uncertainty of the ionization chamber. The motion of the reactor system in the sliding surface (28), (34) is described by the following differential equation

$$\ddot{e}+0.8\dot{e}+0.327e=0, \tag{37}$$

where $e = \hat{Q} - Q_{ref}$ and $\hat{Q}$ is measured value of the actual reactor thermal power Q. Let us assume that

$$\hat{Q} = (1+\psi)Q, \tag{38}$$

where unknown value ψ reflects the uncertainty of the ionization chamber description . It is obvious that if the sliding mode exists in the sliding surface (28), (34) then the tracking error described by the differential equation (37) is vanishing with time, i.e. $\lim e(t) = 0$ as $t \to \infty$. It means that the start up process should be stable. In steady state we will have

$$Q = Q_{ref} - \frac{\psi}{1+\psi}Q_{ref}. \tag{39}$$

The expression (39) tells us that in steady state we should have the stabilization error $\Delta Q = -\dfrac{\psi}{1+\psi}Q_{ref}$.

5.2 The Sliding Mode Controller Design for the Reactor Reference Current Profile Following Mode

First, differentiating payload current tracking error $\varepsilon = I - I_{ref}$ through the system (7)-(26) we can find that the fourth derivative $\varepsilon^{(4)} \sim U$ and $N = 3$ in the equation of the sliding surface (28). It means that the sliding surface (28) actually is the differential equation of the third order

$$\sigma = \dddot{\varepsilon}+\alpha_2\ddot{\varepsilon}+\alpha_1\dot{\varepsilon}+\alpha_0\varepsilon=0 \tag{40}$$

The characteristic equation of the differential equation (28) can be formed as follows:

$$(r+r_1)(r^2 +2\xi\omega_n r+\omega_n^2) = 0, \tag{41}$$

where ω_n, r_1 are to be defined upon requirements to the settling time of the transient process in the sliding mode. This is

$$\omega_n \geq 4/\xi T_s, \qquad r_1 >> \xi\omega_n. \tag{42}$$

The optimal value of ξ in expression (41) is equal to 0.7 [9].

Second, the optimal coefficients of the differential equation (40) have been defined as follows

$$\alpha_0 = r_1\omega_n^2, \quad \alpha_1 = 1.4\omega_n(r_1+\omega_n), \quad \alpha_2 = r_1+1.4\omega_n,. \tag{43}$$

The following value of the settling time has been selected: $T_s = 10 \sec$ and the values of r_1 and ω_n have been specified as follows: $\omega_n = 0.57$, $r_1 = 2$. The coefficients (43) became:

$$\alpha_0 = 0.654, \quad \alpha_1 = 1.927, \quad \alpha_2 = 2.80. \tag{44}$$

Third, the continuous values U^+, U^- in the discontinuous control law (27) were specified as

$$U^+ = 27 \text{ volts}, \quad U^+ = \text{-27 volts} \tag{45}$$

to meet the inequality (30) calculated through the equations of the system (7)-(26).

Remark 5. The sliding mode controller (27), (28), (40), (44), (45) has been designed for the actuator in the form of DC motor, described by the equations (17)-(20). The block-diagram of the sliding mode controller adjusted to a stepper motor is shown at the fig. 4.
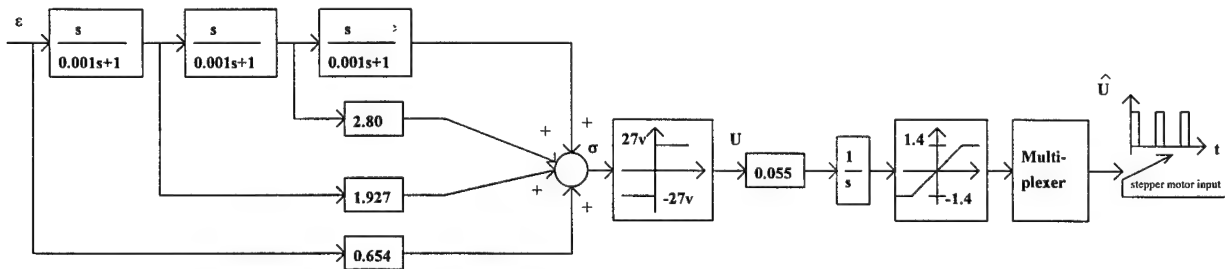
Fig. 4 The block-diagram of the adjusted sliding mode controller
of the reactor reference current profile following mode

6. Simulation of the TOPAZ II Reactor Start Up Mode via Sliding Mode Control

The VisSim simulation package has been used for the simulation of the TOPAZ II reactor system start up mode. The Runge Kutta 4th Order Algorithm with 0.0002 sec step size has been used for the solution of the differential equations during the simulations. The set of the simulations shown at the fig. 5 - 9 corresponds to the absolutely accurate measurement of the thermal power, i.e. $Q_{measured} = Q_{actual}$. The start up controlled mode curves for the cases $Q_{measured} = 0.95 \cdot Q_{actual}$ and $Q_{measured} = 1.05 \cdot Q_{actual}$ are shown at the fig. 10 and 11 correspondingly. . All simulations have been performed for the controlled reactivity range $-4.375 \le \rho_c(t) \le 0.9875$ in \$.

Discussion 1

• Based on the fig. 5 - 9 it is obvious that the desired thermal power profile is tracked almost perfectly during the start up mode. Meanwhile the required angular velocity of a motor doesn't exceed the limit 1.4^0/sec.

• Based on the fig. 10, 11 it is obvious that the thermal power profile following during the start up mode is robust to the uncertainty in the actual thermal power measurement. The steady state tracking error is proportional to the measure of uncertainty ψ. This result is as expected based on the formula (39).

7. Simulation of the TOPAZ II Reactor Payload Current Profile Following via Sliding Mode Control

The VisSim simulation package has been used for the simulation of the TOPAZ II reactor system current profile following mode. The Runge Kutta 4th Order Algorithm with the step size 0.0002 sec has been used during the

33-14

simulations for the solution of the differential equations. The set of simulations, shown at the fig. 12 - 16, has been performed for the controlled reactivity range $-2.875 \leq \rho_c(t) \leq 2.398$ in \$. The next simulation, shown at the fig. 17, has been performed for the controlled reactivity range $-4.375 \leq \rho_c(t) \leq 0.9875$ in \$. The value of the simulation step size has been specified as 0.00025 sec during the second set of the simulations for the controlled reactivity range $-2.875 \leq \rho_c(t) \leq 2.398$ in \$. The results of the simulations with the following backslash characteristic of a motor drive
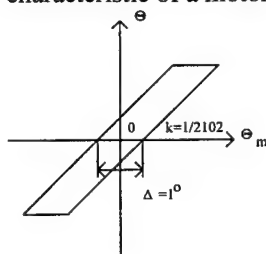
are shown at the fig.18-20.

Discussion 2

• Based on the fig. 12 - 17 we can conclude that the given payload current profile following is very accurate with zero steady state error. This is much better than the requirement to the accuracy of the current stabilization process which is $3.5 \pm 0.35\,A$. Meanwhile the required angular velocity of a motor doesn't exceed the limit 1.4^0/sec.

• Based on the fig. 18 - 20 we can conclude that in the case of backslash the given payload current profile following is also very accurate with zero steady state error.

• On the fig. 19 we can see the oscillations of the control drums. The double magnitude of these oscillations $\approx 0.75^0$ with the angular frequency ≈ 0.2 Hz. To our mind this regime is acceptable for the control drum drive system.

• On the fig. 20 we can see that the angular velocity of the motor position angle changes its sign rapidly. However the limitation (1.4^0/sec) to the value of the angular velocity is met.

 8. Recommendations to the implementation of the sliding mode controllers in TOPAZ II reactor system

Simplified functional diagram of the part of the TOPAZ II Automatic Control System, improved in this work, is depicted at the fig. 21.
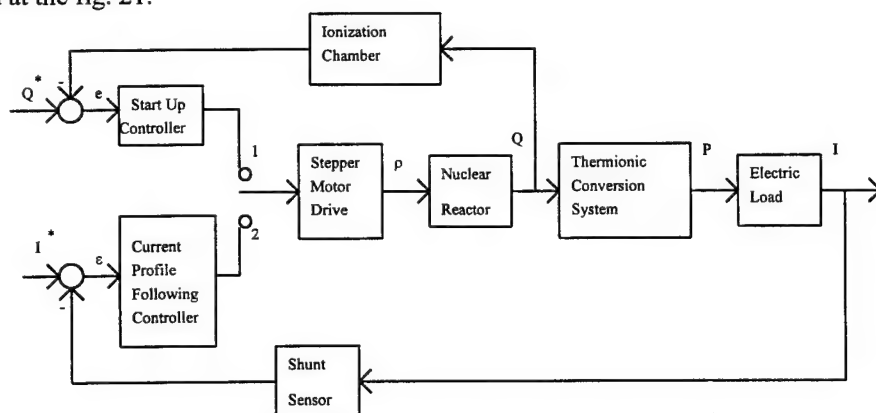
Fig. 21 Simplified functional diagram of the part of the TOPAZ II Automatic Control System

Fig. 5

Desired thermal power profile: start up mode

Fig. 6

Thermal power (Wt): start up mode

Fig. 7

Drum's angle (degrees): start up mode

Fig. 8

Omega (degree/sec)

Temperature (K)

moderator
fuel

Fig. 9

Thermal power (Wt)

Actual thermal power
Reference thermal power

Fig. 10

Thermal power (Wt)

Actual thermal power
Reference thermal power

Fig. 11

Load current (A)

Actual current
Reference current

Fig. 12

33-17
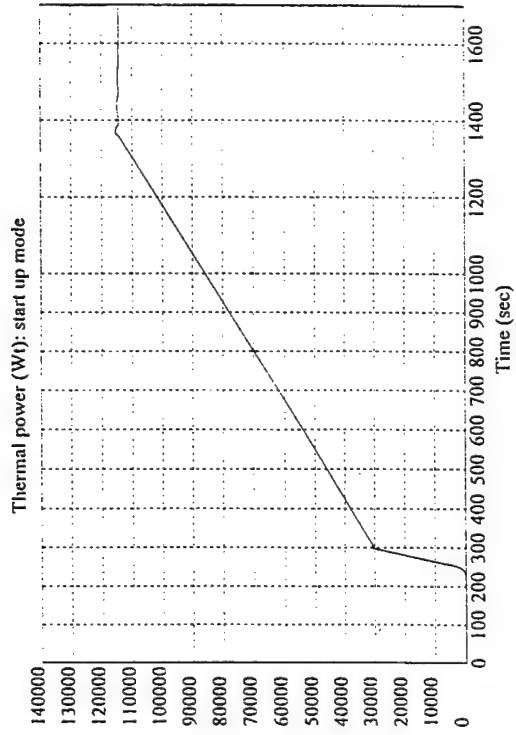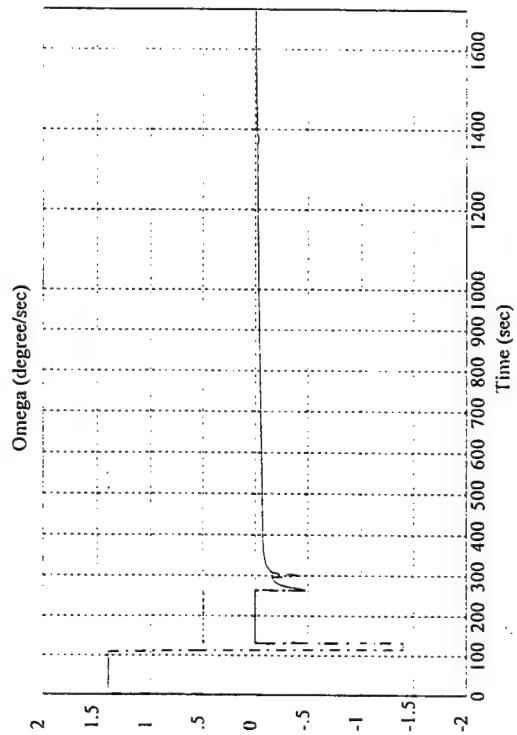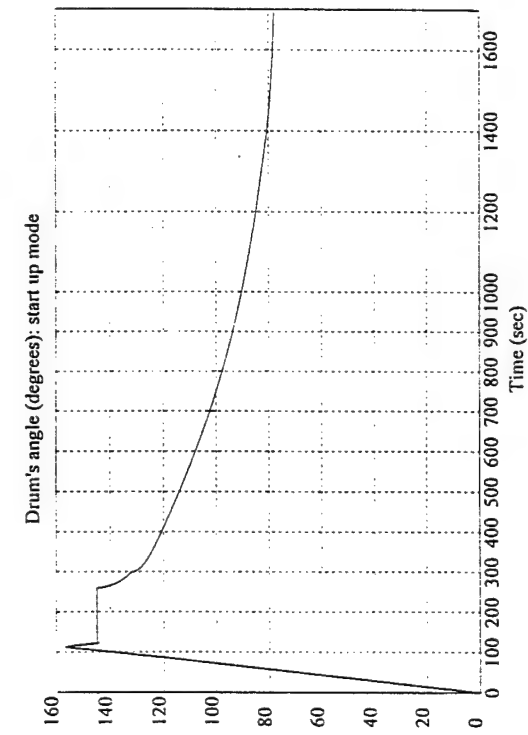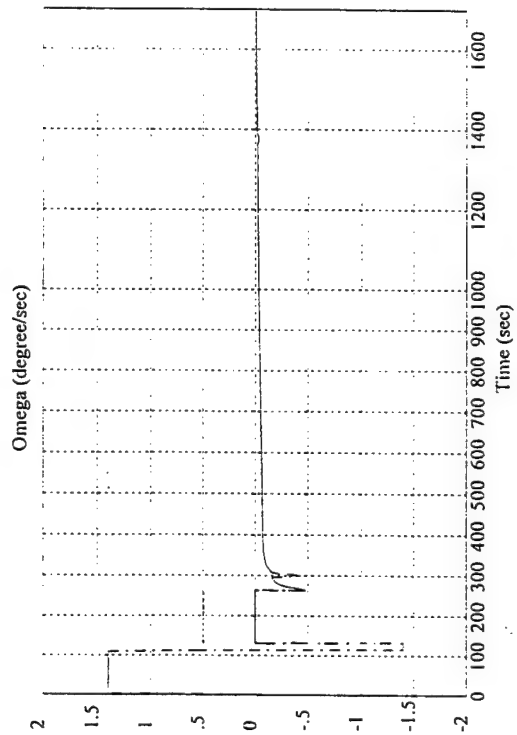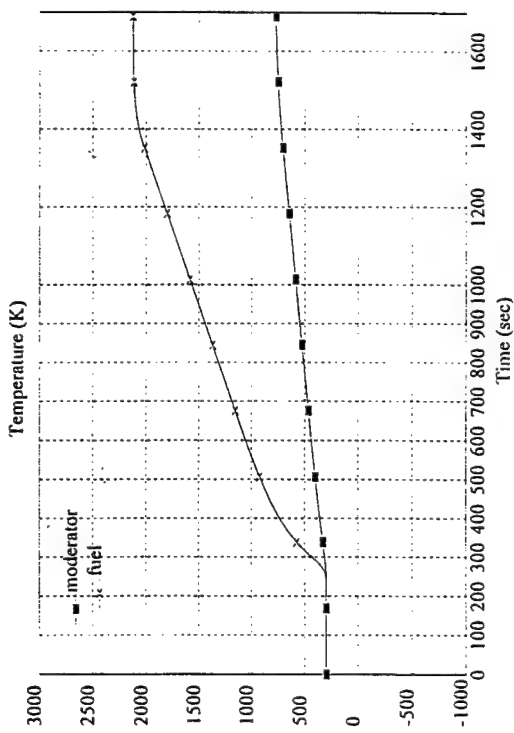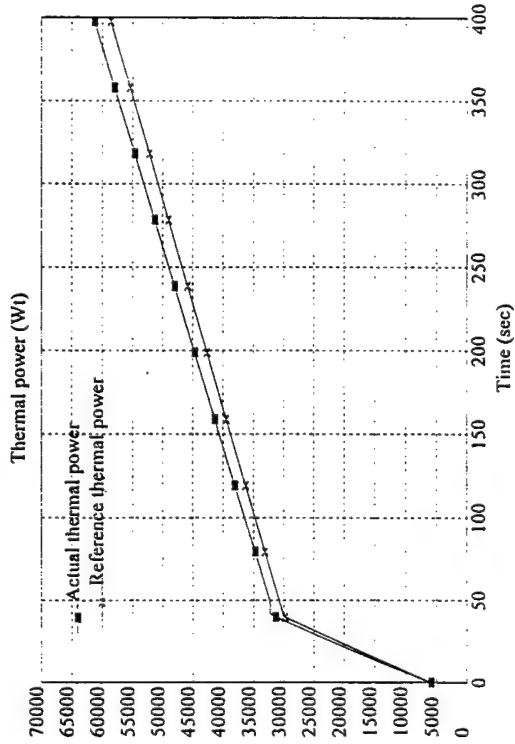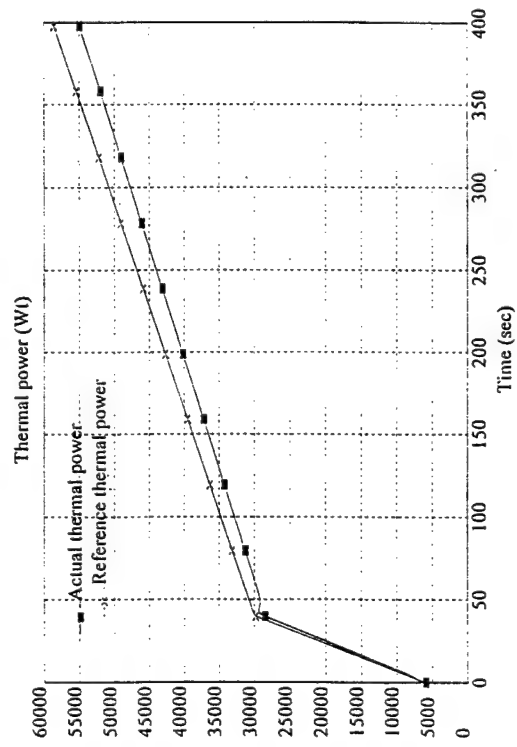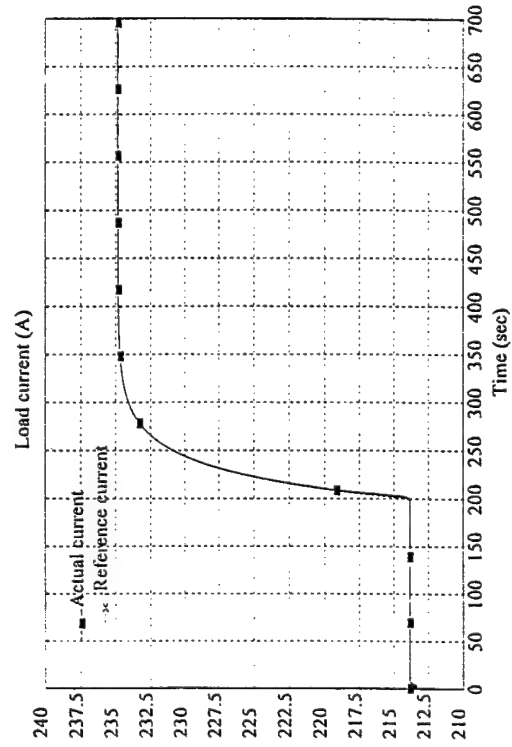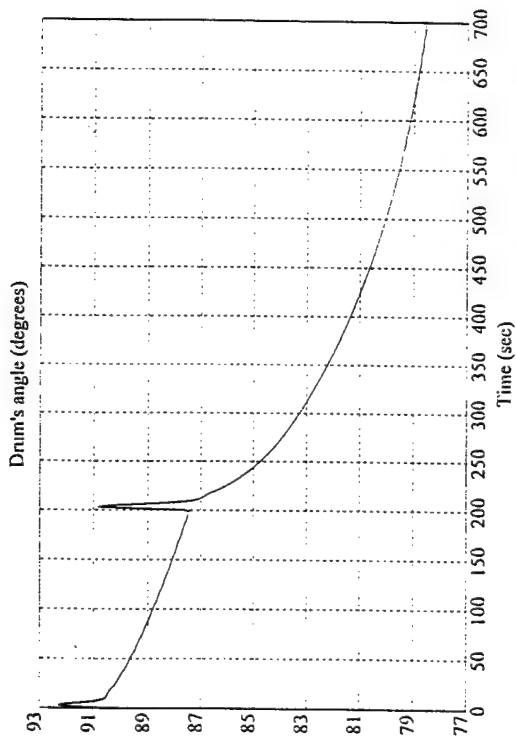
Fig. 13

Fig. 14

Fig. 15

Fig. 16

33-18
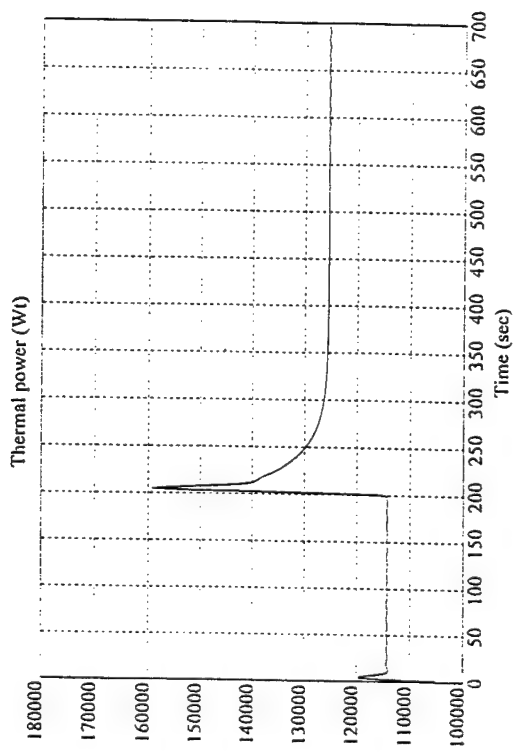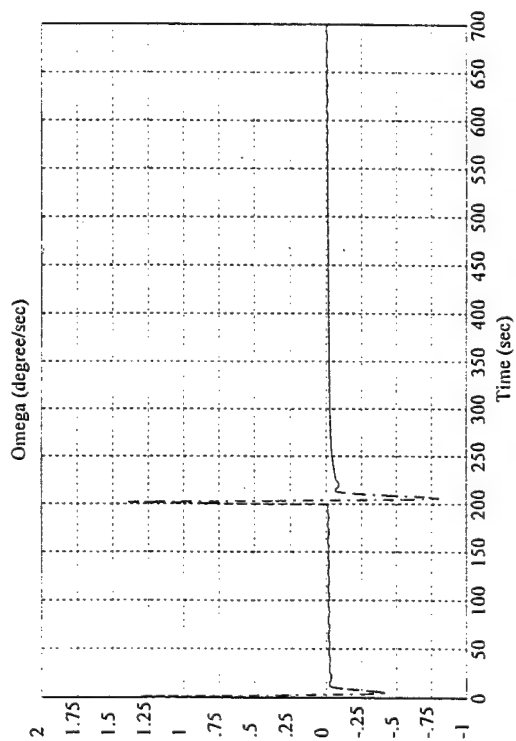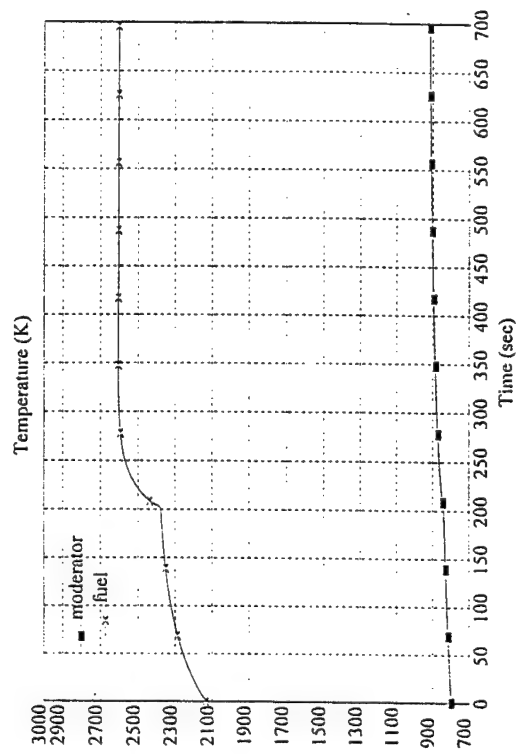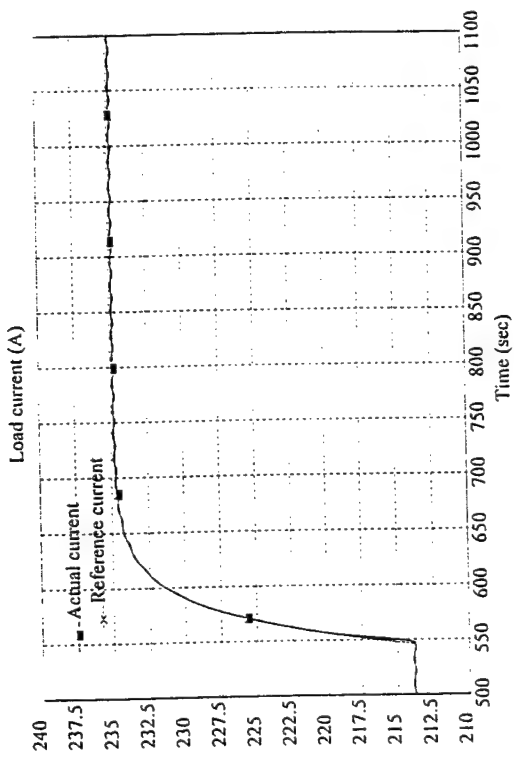
Fig. 17

Fig. 18

Fig. 19

Fig. 20

33-19
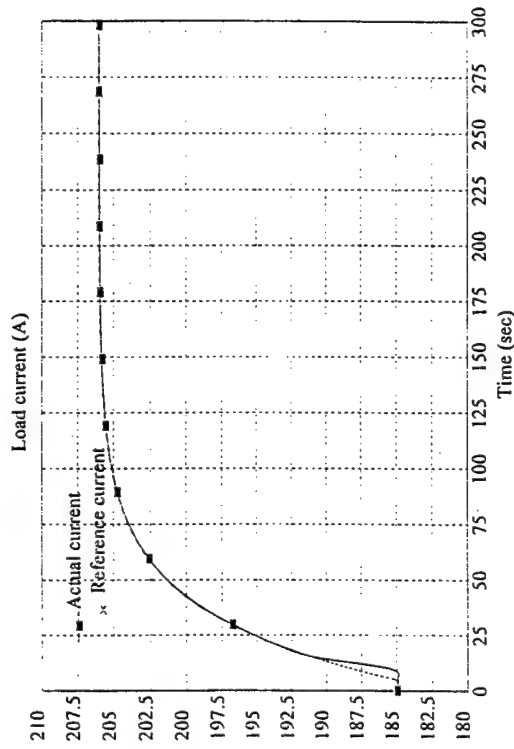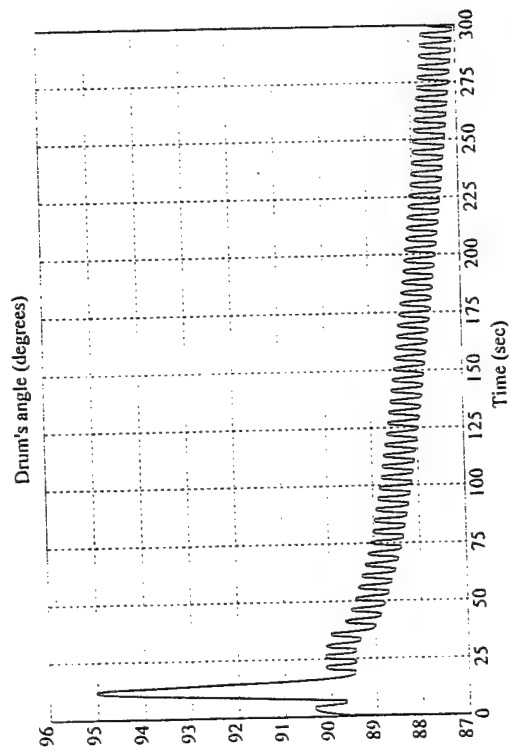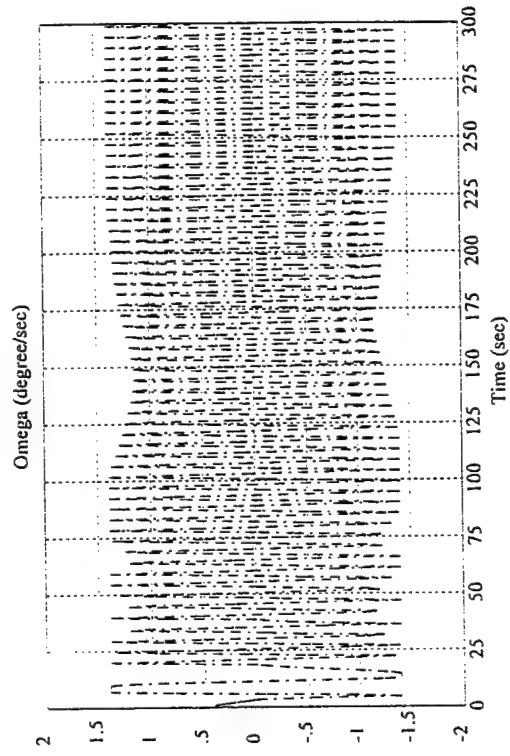
The position 1 at the fig. 21 corresponds to the start up mode and the position 2 corresponds to the reference current profile following mode. Initially the control system has to be turned on the position 1. The control system could be switched in the position 2 after the payload current is ≈ 100 A and output voltage is stabilized on the 27 V level by the mean of the corresponding voltage regulator. Then the reference payload current profile should be followed during the reactor system operation. The other scenario could require the control system to be turned on the position 2 after the start up mode is completely over.

9. Conclusion

The TOPAZ II space nuclear reactor system is considered. The reference thermal power profile following (start up mode) and the reference payload current (electric power) profile following controllers have been designed via sliding mode control technique. The sets of the simulations of the TOPAZ II Reactor System with the sliding mode controllers have been performed. The results of the simulations showed the very accurate reference thermal power profile following and reference payload current profile following which are almost insensitive to the system parameter variations (ionization chamber) and external disturbances.

10. Acknowledgments

11. References

1. Glushkov, E. C. *Personal Communications*, New Mexico Engineering Research Institute, Albuquerque, NM, 1993

2. Gunther, N. G. *Characteristics of the Soviet TOPAZ II Space Power System*, Report SPI-52-1, Space Power Inc., San Jose, CA, 1990.

3. El-Genk, M. S., Xue, H. and Paramonov D. "Start Up Simulation of a Thermionic Space Nuclear Reactor System", *Proc. of the 10th Symp. on Space Nuclear Power Systems*, American Institute of Physics, NY, 1993.

4. Utkin, V. I. *Optimization and Sliding Mode Control*, Springer - Verlag, 1992.

5. DeCarlo, R. A., Zak, S. H. and Matthews, G. P. "Variable structure control of nonlinear multivariable systems: a tutorial," *IEEE Proc.*, Vol. 76, pp. 212-232, 1988.

6. Shtessel, Y. "Tracking Problem Solution in Interconnected Dynamic Systems on Decentralized Sliding Modes," *Modeling, Measurement & Control, B*, AMSE Press, **44**, #3, pp. 53-63, 1992.

7. Kwok, K. S. "Design and Construction of the TOPAZ II Reactor System Real-time Dynamic Simulator", *Proc. of the 11th Symp. on Space Nuclear Power Systems*, American Institute of Physics, NY, 1994.

8. Luppov, A. N., Skorlygin, V. V. and Kwok, K. S. "Control Drum Drive Mechanism and Regulation Characteristics of the TOPAZ II Reactor", *Proc. of the 11th Symp. on Space Nuclear Power Systems*, American Institute of Physics, NY, 1994.

9. Dorf, R. *Modern Control Systems, Six Edition*, Addison-Wesley, 1990.

IMPEDANCES OF COPLANAR CONICAL PLATES IN A UNIFORM DIELECTRIC LENS AND MATCHING CONICAL PLATES FOR FEEDING A PARABOLOIDAL REFLECTOR

Alexander P. Stone
Professor
Department of Mathematics and Statistics

University of New Mexico
Albuquerque, New Mexico 87131

Final Report for:
Summer Faculty Research Program

Phillips Laboratory

September 1994

34-1

IMPEDANCES OF COPLANAR CONICAL PLATES IN A UNIFORM DIELECTRIC LENS AND MATCHING CONICAL PLATES FOR FEEDING A PARABOLOIDAL REFLECTOR

Alexander P. Stone
Professor
Department of Mathematics and Statistics
University of New Mexico

Abstract

In this paper we investigate impedance characteristics of coplanar conical plate geometries which pass through a lens boundary. The plates are initially in a dielectric lens matching to exterior conical plates which serve as a paraboloidal reflector feed. The lens impedance Z_{in}, is a function of the "half-angle", α', of the interior conical plates, but is independent of the ratio, F/D, of focal length to reflector diameter. Various practical choices of this ratio are made and impedances, Z_{out}, of the exterior region are calculated. As $\alpha' \longrightarrow 0$, the ratio Z_{out}/Z_{in} approaches $\sqrt{\epsilon_r}$, where ϵ_r is the relative permittivity of the lens region. As α' increases, the impedance ratio also increases. Numerical and graphical impedance results are presented as a function of F/D, so that one can trade off the various performance characteristics.

IMPEDANCES OF COPLANAR CONICAL PLATES IN A UNIFORM DIELECTRIC LENS AND MATCHING CONICAL PLATES FOR FEEDING A PARABOLOIDAL REFLECTOR

Alexander P. Stone

1 Introduction

One form of an impulse radiating antenna (IRA) consists of a paraboloidal reflector fed by conical transmission lines that propagate a spherical TEM wave, which originates from the focal point of the reflector. A diagram which depicts such an IRA is shown schematically in Figure 1. The design considerations of a uniform dielectric lens useful in launching a spherical
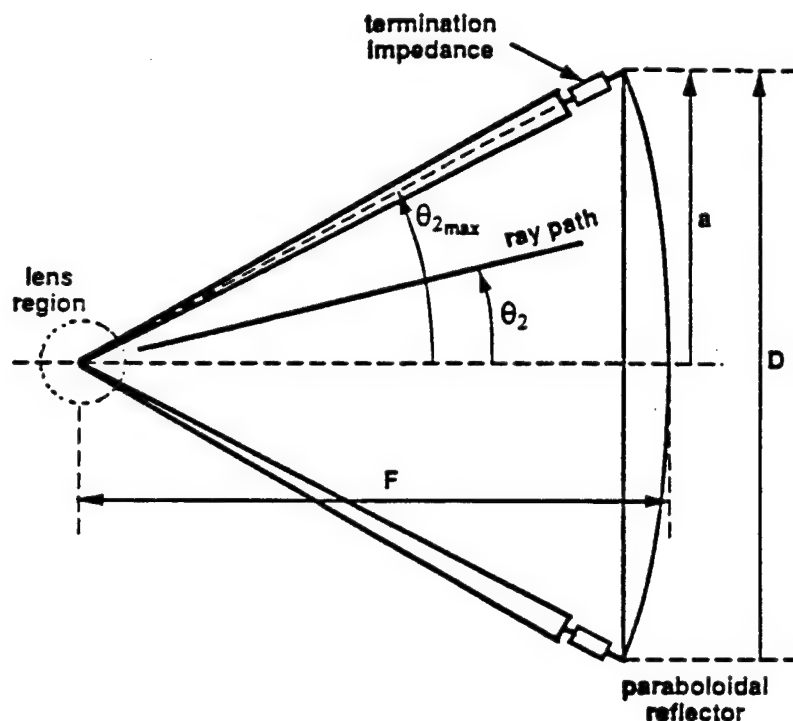
Figure 1: Schematic diagram of Prototype IRA-2

TEM wave onto such a paraboloidal reflector have been investigated in earlier work [1, 2, 3],
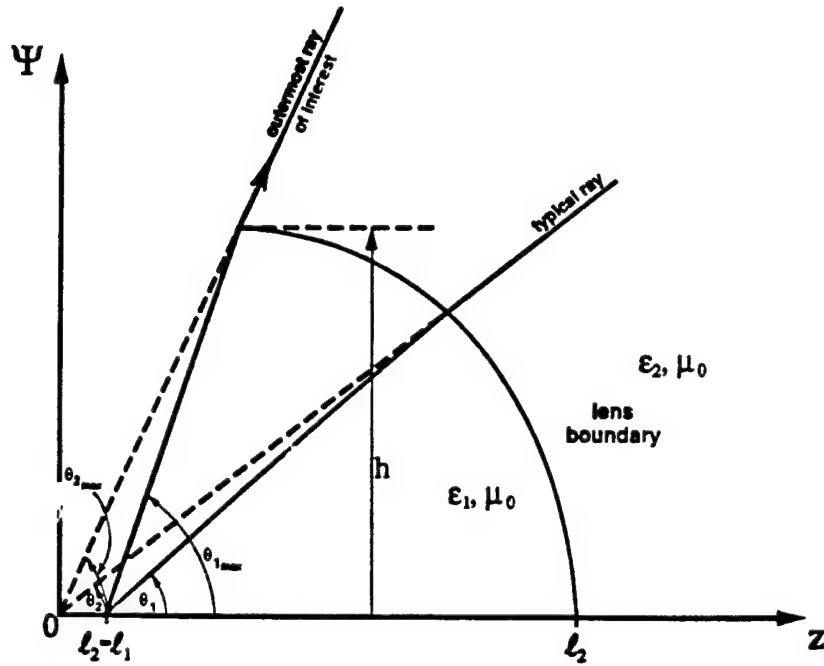
Figure 2: Lens Design Parameters

and the geometrical parameters for the lens design are indicated in Figure 2. The lens design in essence consists of specifying the following parameters.

a) h, the cylindrical radius of the outermost ray which intercepts the reflector,

b) a dielectric medium for the lens, with relative permittivity ϵ_r,

c) the angle, θ_{2max}, from the apex of the conical transmission line (focal point) to the edge of the reflector. This angle is determined by the reflector geometry through the ratio F/D, where F is the focal distance from the center of the reflector and D is the reflector diameter,

d) the angle θ_{1max} made by the line inside the lens with respect to the direction of launch.

There are several reasons for employing dielectric lenses in such an application, and these have been discussed in [3]. The reasons may be briefly summarized as follows. First, the

ideal point source at the focal point is physically realized by a closing switch. The high voltage between the switch electrodes dictates the use of a dielectric medium surrounding the switch enclosure. Secondly, a dielectric medium can also be used to ensure a spherical TEM wave launch inside the lens. All rays emanating from the switch center should exit the lens boundary and end up on a spherical wavefront centered on the focal point at the same instant. The "equal transit-time" condition, which provides the basis for the lens design, generates an optimal spherical wave (not truly TEM at very early times due to different transmission coefficients at different points on the lens surface) launch onto the reflector. Note that θ_{1max} in [1] is taken as 90° and defines the center (symmetry plane) of the conical plate. The frequency dependence of the dielectric constant of the lens medium is also an important consideration in the choice of the material. The dielectric constant chosen is approximately 2.26, which is the dielectric constant for materials such as transformer oil or polyethylene and is approximately frequency independent over the frequency range of interest.

In this report we are concerned with impedance calculations for a lens geometry, which we take as a coplaner conical plate geometry as investigated in [2]. The lens impedance as a function of the half-angle, α', (Figure 3) of the interior conical plates, as well as the impedance outside of the lens, is calculated on the basis of this coplaner conical geometry. The reflector geometry, specified by the F/D ratio, determines the acceptable range for the angle α'. The notation in Figure 3 is the same as that which appears in [2]. The lens characteristic impedance is denoted by Z_{in} while the characteristic impedance of the region outside of the lens is denoted by Z_{out}. These impedances are calculated and tabulated for various F/D ratios in the later sections of this report.
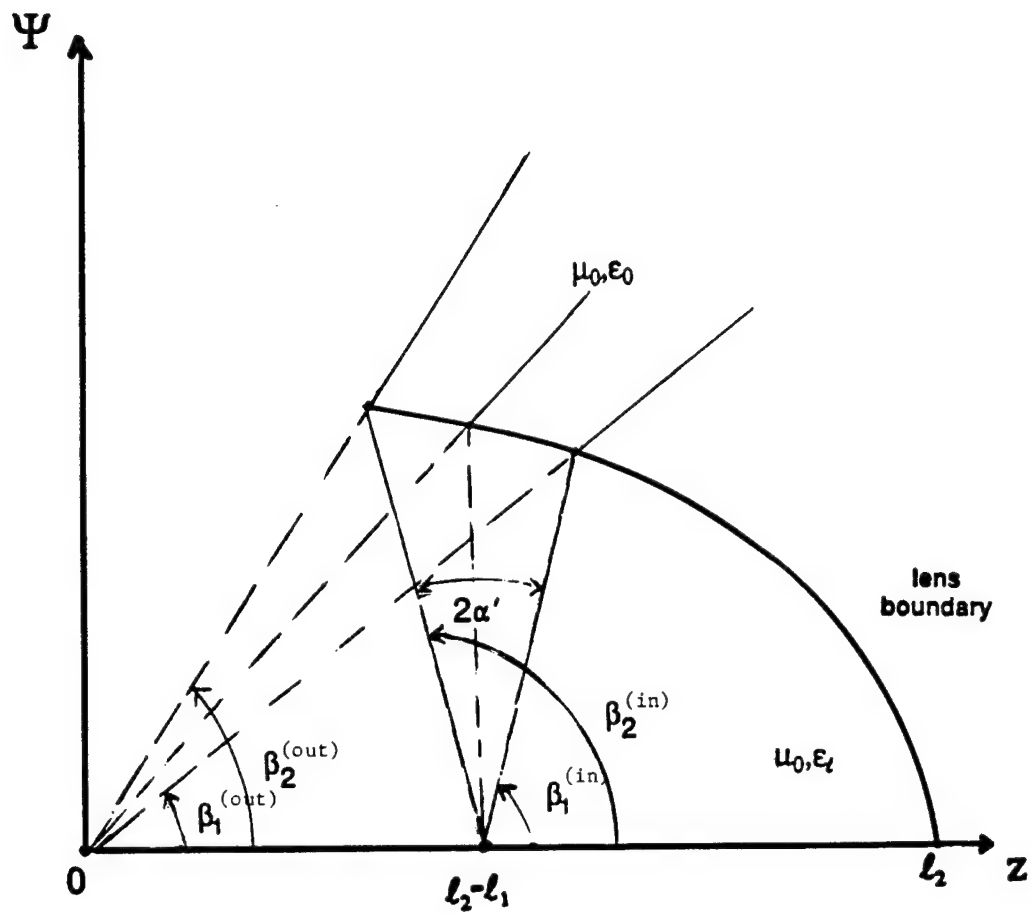
Figure 3: Coplanar Conical Plate Geometry

2 Theoretical Considerations in Lens Design

Let us recall the equal-time condition for a diverging spherical wave in a medium with permittivity ϵ_1 going into another diverging spherical wave in a second medium with permittivity ϵ_2. The equation which describes this equation is derived in [4] and it also appears in [3] in the form

$$\sqrt{\epsilon_1}\left\{[(z_b - \ell_2 + \ell_1)^2 + \Psi_b^2]^{1/2} - \ell_1\right\} = \sqrt{\epsilon_2}\left\{[z_b^2 + \Psi_b^2]^{1/2} - \ell_2\right\} \tag{2.1}$$

where ℓ_1, ℓ_2, z_b, and Ψ_b are as they appear in Figure 2. The coordinates (z_b, Ψ_b) are coordinates of a point on the lens boundary. This equation may then be used to obtain an analytical expression for the lens boundary curve, which expresses the angle θ_2 as a function of θ_1. The resulting expression is

$$\cos(\theta_2) = \frac{AB\sin^2(\theta_1) + |B\cos(\theta_1) - A\epsilon_r^{\frac{1}{2}}|\sqrt{\left[B^2 - 2AB\epsilon_r^{\frac{1}{2}}\cos(\theta_1) + A^2\epsilon_r\right] - A^2\sin^2(\theta_1)}}{B^2 - 2AB\epsilon_r^{\frac{1}{2}}\cos(\theta_1) + A^2\epsilon_r}$$

$$\tag{2.2}$$

where

$$A = (\ell_2/\ell_1) - 1$$
$$B = (\ell_2/\ell_1) - \epsilon_r^{\frac{1}{2}}$$

The lens boundary curve is specified by first choosing an appropriate F/D ratio which determines the angle θ_{2max} of Figure 1. The lens parameters h, ϵ_r, and θ_{1max} are then chosen (consistent with certain constraints which are discussed in [1]) so that the lens parameters ℓ_1/h and ℓ_2/h are obtained and a boundary curve is generated with $0 \leq \theta_1 \leq \theta_{1max}$ and $0 \leq \theta_2 \leq \theta_{2max}$. Note that θ_{2max} is not the largest θ_2 on the lens boundary, but that θ_2 which corresponds to the edge of the reflector and similarly for θ_{1max}. In the cases which we investigate the angle θ_{1max} will be 90°. For the chosen value of F/D, a family of lens designs may then be generated by various choices of θ_{1max}, subject to the constraints mentioned in

[1]. In Figure 4, lens boundary curves for $\theta_{1max} = 90°$ and F/D values of 0.25, 0.3, 0.33, 0.4, and 0.5 are displayed. Note that larger values of F/D lead to larger lenses, while the value of θ_{2max} decreases with increasing F/D. The limiting or trivial case occurs when $F/D = 0.25$. In this case, unlike the cases where $F/D > 0.25$, reflections from the lens surface (spherical) do occur, bu the TEM wave is not distorted. There are also some limitations on how small θ_2 can be for a given θ_{1max}. These are based on the slope of the lens boundary matching the ray direction.
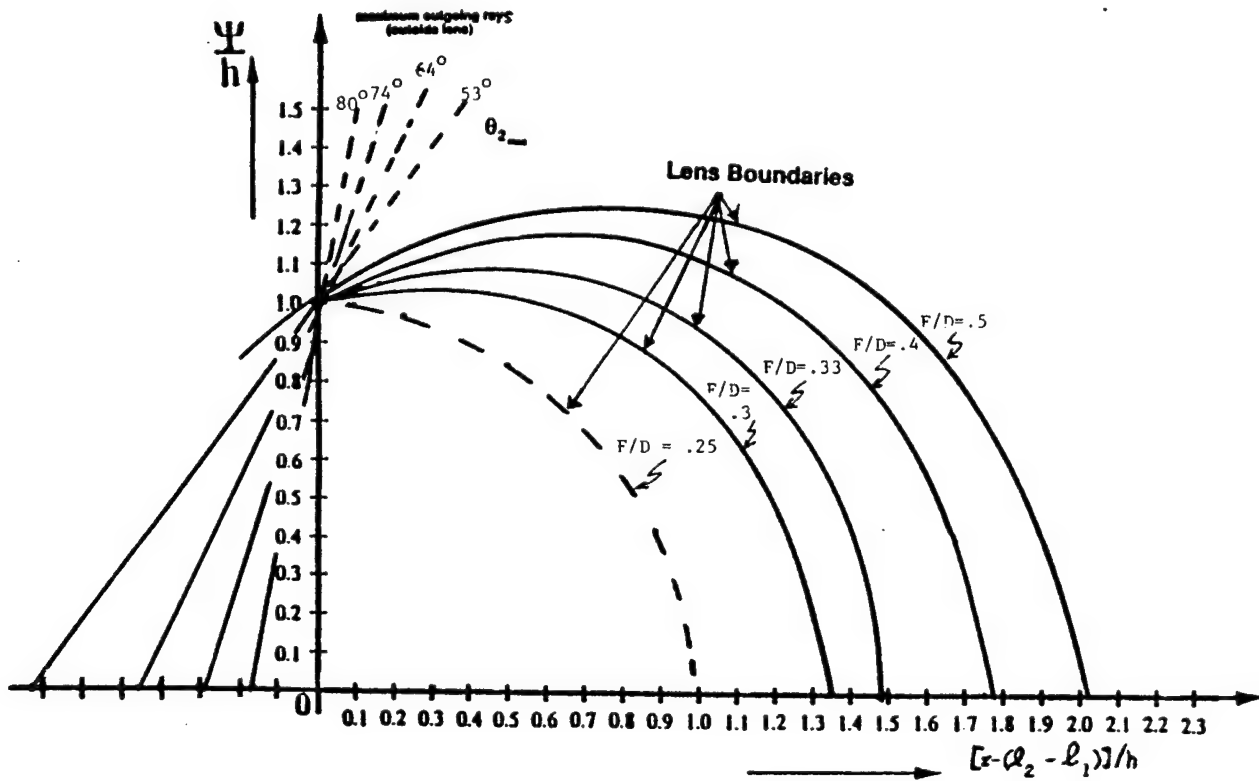
Figure 4: Lens Shapes with $\theta_{1max} = 90°$, $\epsilon_r = 2.26$ (Cylindrical Coordinates)

3 Impedance Calculations

We begin by fixing the F/D ratio which determines the angle θ_{2max} of Figure 1. Since the angle θ_{1max} is 90° for all chosen values of F/D, and $\epsilon_r = 2.26$, the range of the angle α' is determined. Thus for all values of F/D considered, the value of Z_{in} depends only on α'. The lens characteristic impedance Z_{in} can then be calculated by the procedure described in this section. Note that $2\alpha' = \beta_2^{(in)} - \beta_1^{(in)}$ and so the value of α' determines the angles $\beta_i^{(in)}, i = 1, 2$. The angles $\beta_i^{(out)}$ are then found from equation (2.2) and thus the lens shape will be determined.

The problem of calculating the characteristic impedance of the lens region and the region exterior to the lens can now be considered. The geometry, that of coplanar conical plates, is described in Figure 3. A formula for the characteristic impedance of such a geometry appears in [2] and is given by

$$Z_{in} = \frac{Z_0}{\sqrt{\epsilon_r}} \frac{K(m_{in})}{K'(m_{in})} \tag{3.1}$$

where ϵ_r is the relative permittivity for the lens material and the parameter m_{in} is given by

$$m_{in} = \frac{\tan^2(\beta_1^{(in)}/2)}{\tan^2(\beta_2^{(in)}/2)} \tag{3.2}$$

The impedance Z_0 is 376.73 ohms. The quantities $K(m_{in})$ and $K'(m_{in})$ are complete elliptic integrals of the first kind. Formulas for these integrals appear in many places (for example, in [5]) and are given by

$$K(m_{in}) = \int_0^{\pi/2} (1 - m_{in}\sin^2(\theta))^{-\frac{1}{2}} d\theta \tag{3.3}$$

$$K'(m_{in}) = K(1 - m_{in}) \equiv K(m_1^{(in)}) \tag{3.4}$$

where $0 \leq m_{in} < 1$ and $m_1^{(in)} = 1 - m_{in}$. We may also introduce a geometric impedance factor $f_g^{(in)}$ as

$$f_g^{(in)} = \frac{K(m_{in})}{K'(m_{in})} \tag{3.5}$$

34-9

and thus rewrite (3.1) in the form

$$Z_{in} = f_g^{(in)} Z_0 / \sqrt{\epsilon_r} \tag{3.6}$$

Similar expressions hold for the region exterior to the lens. Thus we take for the characteristic impedance of this region

$$Z_{out} = Z_0 \frac{K(m_{out})}{K'(m_{out})} \tag{3.7}$$

where

$$m_{out} \equiv \frac{\tan^2(\beta_1^{(out)}/2)}{\tan^2(\beta_2^{(out)}/2)} \tag{3.8}$$

and

$$K'(m_{out}) = K(1 - m_{out}) \tag{3.9}$$

Likewise a geometric factor, $f_g^{(out)}$, is given by

$$f_g^{(out)} = \frac{K(m_{out})}{K'(m_{out})}. \tag{3.10}$$

The elliptic integrals which depend on the parameters m_{in} and m_{out} may be evaluated by the tables in [5] or by use of the Mathematica program. The integrals can also be evaluated numerically by expanding the integrand and performing a term-wise integration which yields, for $|m_{in}| < 1$,

$$K(m_{in}) = \frac{\pi}{2} \left\{ 1 + \left(\frac{1}{2}\right)^2 m_{in} + \left(\frac{1.3}{2.4}\right)^2 m_{in}^2 + \dots \right\} \tag{3.11}$$

$$= \frac{\pi}{2} \sum_{n=0}^{\infty} \frac{1}{4^n} \binom{2n}{n}^2 m_{in}^n \tag{3.12}$$

where

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} .$$

When m_{in}, or m_{out}, is near 0 or 1, there are expressions for $K(m)/K'(m)$ which may be derived. We consider approximations for m_{in} first.

The formula (3.2) for m_{in} may be reexpressed in the form

$$m_{in} = \frac{(1 - \cos(\beta_1^{(in)})) \, (1 + \cos(\beta_2^{(in)}))}{(1 + \cos(\beta_1^{(in)})) \, (1 - \cos(\beta_2^{(in)}))} \tag{3.13}$$

and since

$$\beta_1^{(in)} = \frac{\pi}{2} - \alpha'$$
$$\beta_2^{(in)} = \frac{\pi}{2} + \alpha' \tag{3.14}$$

we then obtain, for small α',

$$m_{in} \sim \frac{(1 - \sin(\alpha))^2}{(1 + \sin(\alpha))^2} \sim \left(\frac{1 - \alpha}{1 + \alpha}\right)^2 . \tag{3.15}$$

Hence

$$m_{in} \sim 1 - 4\alpha' \, , \text{ as } \alpha \longrightarrow 0 \tag{3.16}$$

and so

$$m_1^{(in)} = 1 - m_{in} \sim 4\alpha' \, , \text{ as } \alpha \longrightarrow 0 \, . \tag{3.17}$$

We also have, from [5],

$$\frac{K'(m_{in})}{K(m_{in})} \sim \frac{1}{\pi} ln\left(\frac{16}{m_{in}}\right) \, , \text{ as } m_{in} \longrightarrow 0 \tag{3.18}$$

$$\frac{K(m_{in})}{K'(m_{in})} \sim \frac{1}{\pi} ln\left(\frac{16}{1 - m_{in}}\right) \, , \text{ as } m_{in} \longrightarrow 1 \tag{3.19}$$

Similar expressions for m_{out} may also be derived from

$$m_{out} = \frac{(1 - \cos(\beta_1^{(out)})) \, (1 + \cos(\beta_2^{(out)}))}{(1 + \cos(\beta_1^{(out)})) \, (1 - \cos(\beta_2^{(out)}))} . \tag{3.20}$$

If we use formula (2.2) and let $\alpha' \longrightarrow 0$ we find

$$m_{out} \sim \frac{B^2 + 2AB\epsilon_r^{\frac{1}{2}}\alpha' + A^2\epsilon_r}{B^2 - 2AB\epsilon_r^{\frac{1}{2}}\alpha' + A^2\epsilon_r} . \tag{3.21}$$

Hence, if $m_{in} \longrightarrow 1$, $\alpha' \longrightarrow (1 - m_{in})/4$ as $\alpha' \longrightarrow 0$.

Thus

$$m_{out} \sim \frac{1 + a(1 - m_{in})}{1 - a(1 - m_{in})} \, , \text{ as } m_{in} \longrightarrow 1 \, , \tag{3.22}$$

34-11

where

$$a = \frac{AB\epsilon_r^{\frac{1}{2}}}{2(B^2 + A^2\epsilon_r)} .$$

Thus if $m_{in} \longrightarrow 1$, the impedance ratio

$$\frac{Z_{out}}{Z_{in}} = \sqrt{\epsilon_r} \frac{K(m_{out})}{K'(m_{out})} \cdot \frac{K'(m_{in})}{K(m_{in})}$$

$$\sim \sqrt{\epsilon_r} \frac{\ell n \frac{16}{[1-(1+a(1-m_{in}))^2]}}{\ell n \frac{16}{1-m_{in}}} \longrightarrow \epsilon_r^{\frac{1}{2}} \qquad (3.23)$$

as $m_{in} \longrightarrow 1$.

Armed with the above formulas we may now calculate the impedances Z_{in} and Z_{out}, and also the ratio Z_{out}/Z_{in} for various practical values of F/D. In our impedance calculations we have chosen F/D ratios of 2.5, 3.0, 3.3, 4.0 and 5.0. For each of these F/D ratios, tabulations of α', $\beta_i^{(in)}$, m_{in}, $f_g^{(in)}$, Z_{in}, $\beta_i^{(out)}$, m_{out}, $f_g^{(out)}$, Z_{out} and Z_{out}/Z_{in} are given. The data appears in Tables 1 through 5, while graphical results are presented in Figure 5, which gives plots of Z_{out}/Z_{in} versus α', for each of the chosen values of F/D. The special case where $F/D = 0.25$ was mentioned in Section 2. Note that for our choice of $\epsilon_r = 2.26$ and for the limiting value of $F/D = 0.25$, we have $Z_{out}/Z_{in} = \sqrt{\epsilon_r} = 1.503$. Also, on each of the Tables 2 through 5 the calculation for $Z_{out} = 400$ ohms is listed separately.

4   Conclusions

In Tables 2 through 5 we observe that as α' increases, the ratio Z_{out}/Z_{in} increases. Small values of α' correspond to larger values of Z_{in}, and so as $\alpha' \longrightarrow 0$, Z_{in} increases and smaller values of Z_{out}/Z_{in} imply a better match at the lower frequencies (radian wavelengths of the order of h or larger) for which a transmission-line approximation is appropriate. As $\alpha' \longrightarrow 0$, $Z_{out}/Z_{in} \longrightarrow \epsilon_r^{\frac{1}{2}} = 1.503$, which is the ratio of the wave impedances of the media. For a ray propagating along the z axis (in the positive z direction) this represents a
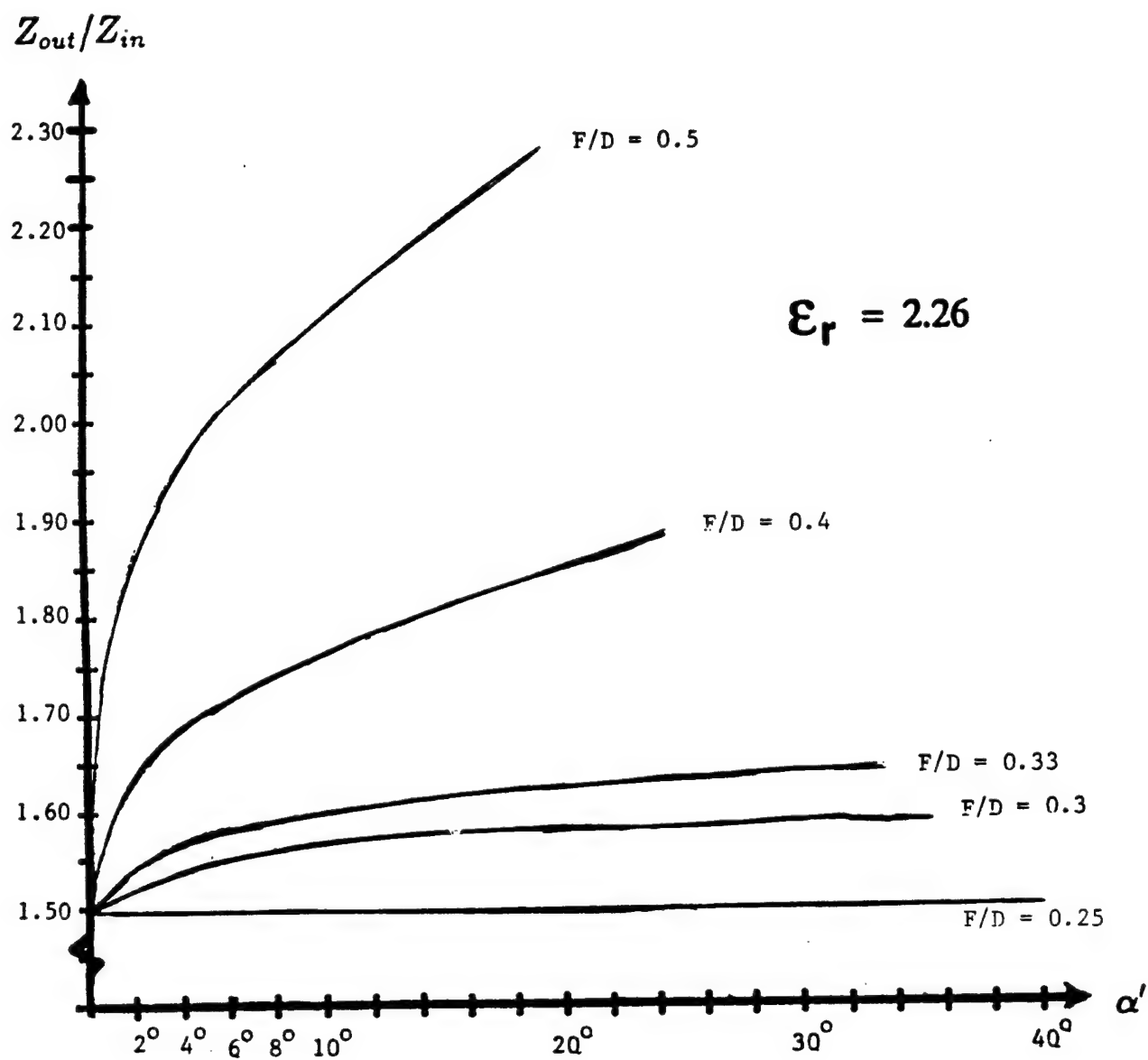
Figure 5: Impedance Ratio, Z_{out}/Z_{in}, versus half-angle, α'

transmission coefficient of

$$T_t = \frac{2\epsilon_r^{\frac{1}{2}}}{\epsilon_r^{\frac{1}{2}} + 1} = 1.201 \qquad (4.24)$$

while the smaller values of α' make the transmission-line transmission coefficient

$$T_Z = \frac{2\frac{Z_{out}}{Z_{in}}}{\frac{Z_{out}}{Z_{in}} + 1} \qquad (4.25)$$

approach the same value. Note that for $F/D = 0.25$ the two transmission coefficients are the same for all α'.

The special case of $F/D = 0.25$ (a spherical lens) has the property that waves transmitted and reflected via the lens boundary are spherical TEM waves. However the reflected wave then is focused on the apex of the interior conical transmission line where it reflects back toward and through the lens boundary (with a fast rise time) unless there is a matched load at this apex. Larger F/D values make this wave reflected from the lens boundary dispersed in the lens and exiting the lens giving a smoother transition from the initially-transmitted fast-rising wave to the late-time behavior. It would then appear that F/D should be larger than 0.25, but not excessively so, since lens size increases with increasing F/D.

α' (degrees)	$\beta_1^{(in)}$ (degrees)	$\beta_2^{(in)}$ (degrees)	m_{in}	$f_g^{(in)}$	Z_{in} (ohms)
0.5	89.5	90.5	.9657	1.951	488.8
1.0	89	91	.9326	1.730	433.5
2.0	88	92	.8697	1.509	378.2
3.0	87	93	.8110	1.380	345.8
4.0	86	94	.7562	1.288	322.9
5.0	85	95	.7050	1.217	305.1
6.0	84	96	.6573	1.159	290.5
7.0	83	97	.6127	1.110	278.2
8.0	82	98	.5710	1.067	267.5
9.0	81	99	.5321	1.030	258.1
10.0	80	100	.4957	0.9961	249.6
11.0	79	101	.4618	0.9656	242.0
12.0	78	102	.4300	0.9377	235.0
15.0	75	105	.3467	0.8660	217.0
18.0	72	108	.2786	0.8072	202.3
21.0	69	111	.2231	0.7571	189.7
24	66	114	.1779	0.7135	178.8
27	63	117	.1410	0.6747	169.1
30	60	120	.1111	0.6396	160.3
33	57	123	.0869	0.6076	152.3
36	54	126	.0674	0.5780	144.9
42	48	132	.0393	0.5245	131.4
48	42	138	.0217	0.4766	119.4
54	36	144	.0111	0.4325	108.4
60	30	150	.0052	0.3909	97.9
66	24	156	.0020	0.3504	87.8

Table 1: Coplanar Conical Plate Parameters Inside Lens (independent of F/D)

α' (degrees)	$\beta_1^{(out)}$ (degrees)	$\beta_2^{(out)}$ (degrees)	m_{out}	$f_g^{(out)}$	Z_{out} (ohms)	$\dfrac{Z_{out}}{Z_{in}}$
0.5	79.12	80.03	.9684	1.977	744.7	1.524
1.0	78.67	80.48	.9378	1.756	661.6	1.526
2.0	77.77	81.39	.8794	1.536	578.5	1.530
3.0	76.87	82.30	.8246	1.406	529.8	1.532
4.0	75.96	83.21	.7732	1.315	495.3	1.534
5.0	75.06	84.12	.7249	1.244	468.5	1.536
6.0	74.17	85.03	.6796	1.186	446.6	1.537
7.0	73.27	85.95	.6370	1.136	428.1	1.539
8.0	72.37	86.86	.5970	1.094	412.0	1.540
9.0	71.47	87.78	.5594	1.056	397.8	1.542
10.0	70.58	88.71	.5241	1.022	385.1	1.543
11.0	69.68	89.62	.4909	0.9918	373.6	1.544
12.0	68.79	90.55	.4598	0.9638	363.1	1.545
15.0	66.11	93.33	.3770	0.8920	336.0	1.548
18.0	63.44	96.14	.3082	0.8330	313.8	1.551
21.0	60.77	98.97	.2511	0.7827	294.9	1.554
24.0	58.11	101.8	.2037	0.7388	278.3	1.557
27.0	55.46	104.7	.1644	0.6997	263.6	1.559
30.0	52.80	107.6	.1319	0.6644	250.3	1.561
33.0	50.15	110.6	.1050	0.6320	238.1	1.564
36.0	47.50	110.8	.0920	0.6147	231.6	1.599
42.0	42.21	104.3	.9000	0.6120	230.5	1.754
48.0	36.92	96.82	.0878	0.6089	229.4	1.920
54.0	31.64	88.22	.0855	0.6056	228.1	2.105
60.0	26.37	78.19	.0831	0.6021	226.8	2.316
66.0	21.09	66.44	.0808	0.5988	225.6	2.569

Note: When $\alpha' = 8.8°$, $Z_{in} = 259.7\Omega$, $Z_{out} = 400\Omega$, and $Z_{out}/Z_{in} = 1.542$

Table 2: Coplanar Conical Plate Parameters Outside Lens ($F/D = 0.3$)

α' (degrees)	$\beta_1^{(out)}$ (degrees)	$\beta_2^{(out)}$ (degrees)	m_{out}	$f_g^{(out)}$	Z_{out} (ohms)	$\dfrac{Z_{out}}{Z_{in}}$
0.5	73.90	74.70	.9715	2.011	757.6	1.550
1.0	73.50	75.09	.9439	1.790	674.4	1.556
2.0	72.70	75.89	.8909	1.570	591.3	1.564
3.0	71.91	76.68	.8409	1.440	542.7	1.569
4.0	71.11	77.48	.7936	1.349	508.2	1.574
5.0	70.31	78.27	.7490	1.278	481.4	1.578
6.0	69.51	79.06	.7068	1.220	459.5	1.581
7.0	68.70	79.85	.6670	1.170	441.0	1.585
8.0	67.90	80.64	.6293	1.128	424.9	1.588
9.0	67.10	81.43	.5937	1.090	410.7	1.592
10.0	66.29	82.22	.5600	1.057	398.1	1.595
11.0	65.49	83.01	.5282	1.026	386.6	1.597
12.0	64.68	83.79	.4981	0.9982	376.1	1.600
15.0	62.26	86.15	.4172	0.9266	349.1	1.608
18.0	59.82	88.50	.3488	0.8679	326.9	1.616
21.0	57.39	90.84	.2909	0.8179	308.1	1.624
24.0	54.94	93.18	.2419	0.7744	291.7	1.632
27.0	52.49	95.51	.2005	0.7357	277.2	1.639
30.0	50.03	97.84	.1654	0.7008	264.0	1.647
33.0	47.57	100.2	.1359	0.6689	252.0	1.655
36.0	45.10	102.5	.1110	0.6395	240.9	1.663

Note: When $\alpha' = 9.8°$, $Z_{in} = 250.9\Omega$, $Z_{out} = 400\Omega$, and $Z_{out}/Z_{in} = 1.594$

Table 3: Coplanar Conical Plate Parameters Outside Lens ($F/D = 0.33$)

α' (degrees)	$\beta_1^{(out)}$ (degrees)	$\beta_2^{(out)}$ (degrees)	m_{out}	$f_g^{(out)}$	Z_{out} (ohms)	$\dfrac{Z_{out}}{Z_{in}}$
0.5	63.77	64.29	.9799	2.123	799.8	1.636
1.0	63.50	64.55	.9602	1.902	716.7	1.653
2.0	62.97	65.06	.9220	1.682	633.6	1.675
3.0	62.42	65.56	.8853	1.553	585.0	1.691
4.0	61.87	66.05	.8501	1.461	551.0	1.705
5.0	61.31	66.53	.8164	1.390	523.8	1.717
6.0	60.74	67.00	.7840	1.332	502.0	1.728
7.0	60.17	67.46	.7529	1.284	483.6	1.738
8.0	59.58	67.90	.7231	1.241	467.6	1.748
9.0	58.99	68.34	.6946	1.204	453.6	1.758
10.0	58.40	68.76	.6672	1.171	441.1	1.767
11.0	57.79	69.17	.6409	1.141	429.7	1.776
12.0	57.19	69.57	.6158	1.113	419.4	1.785
15.0	55.32	70.68	.5465	1.043	393.1	1.811
18.0	53.41	71.65	.4855	0.9869	371.8	1.838
21.0	51.45	72.48	.4321	0.9395	353.9	1.865
24.0	49.45	73.14	.3852	0.8990	338.7	1.894
27.0	47.41	73.63	.3441	0.8639	325.4	1.925
30.0	45.34	73.91	.3082	0.8329	313.8	1.958
33.0	43.23	73.97	.2768	0.8056	303.5	1.993
36.0	41.10	73.79	.2494	0.7812	294.3	2.032
42.0	36.76	72.58	.2047	0.7398	278.7	2.120
48.0	32.33	70.06	.1710	0.7065	266.1	2.228
54.0	27.83	65.98	.1457	0.6798	256.1	2.363
60.0	23.27	60.07	.1269	0.6586	248.1	2.533
66.0	18.67	52.09	.1132	0.6422	241.9	2.755

Note: When $\alpha' = 14.2°$, $Z_{in} = 221.7\Omega$, $Z_{out} = 400\Omega$, and $Z_{out}/Z_{in} = 1.804$

Table 4: Coplanar Conical Plate Parameters Outside Lens ($F/D = 0.4$)

α' (degrees)	$\beta_1^{(out)}$ (degrees)	$\beta_2^{(out)}$ (degrees)	m_{out}	$f_g^{(out)}$	Z_{out} (ohms)	$\dfrac{Z_{out}}{Z_{in}}$
0.5	53.01	53.24	.9899	2.347	884.0	1.809
1.0	52.89	53.35	.9800	2.126	800.9	1.848
2.0	52.64	53.56	.9606	1.905	717.8	1.898
3.0	52.37	53.75	.9415	1.777	669.3	1.935
4.0	52.08	53.92	.9228	1.685	634.9	1.966
5.0	51.79	54.08	.9046	1.615	608.3	1.994
6.0	51.48	54.22	.8868	1.557	586.6	2.019
7.0	51.15	54.34	.8694	1.509	568.3	2.043
8.0	50.81	54.44	.8525	1.467	552.6	2.066
9.0	50.46	54.53	.8360	1.430	538.7	2.088
10.0	50.09	54.59	.8199	1.397	526.4	2.109
11.0	49.72	54.64	.8043	1.368	515.3	2.129
12.0	49.33	54.67	.7891	1.341	505.2	2.150
15.0	48.09	54.64	.7460	1.273	479.7	2.210
18.0	46.76	54.43	.7066	1.219	459.4	2.271
21.0	45.33	54.03	.6708	1.175	442.7	2.333
24.0	43.83	53.45	.6383	1.138	428.7	2.398
27.0	42.25	52.67	.6090	1.106	416.8	2.465
30.0	40.60	51.70	.5827	1.079	406.5	2.536
33.0	38.88	50.54	.5590	1.056	397.7	2.612
36.0	37.11	49.19	.5379	1.035	390.0	2.692
42.0	33.43	45.93	.5021	1.002	377.5	2.872
48.0	29.57	41.95	.4738	0.9764	367.8	3.080
54.0	25.57	37.32	.4517	0.9567	360.4	3.325
60.0	21.47	32.08	.4346	0.9417	354.8	3.622
66.0	17.27	26.33	.4217	0.9305	350.5	3.992

Note: When $\alpha' = 32.2°$, $Z_{in} = 154.3\Omega$, $Z_{out} = 400\Omega$, and $Z_{out}/Z_{in} = 2.592$

Table 5: Coplanar Conical Plate Parameters Outside Lens ($F/D = 0.5$)

References

[1] C. E. Baum, J. J. Sadler, and A. P. Stone, A Uniform Dielectric Lens for Launching a Spherical Wave into a Paraboloidal Reflector, Sensor and Simulation Note 360, July 1993.

[2] E. G. Farr and C. E. Baum, Prepulse Associated with the TEM Feed of an Impulse Radiating Antenna, Sensor and Simulation Note 337, March 1992.

[3] D. V. Giri, Design Considerations of a Uniform Dielectric Lens for Launching a Spherical TEM Wave on the the Protype IRA, Protype IRA Memos, Memo 3, May 1994.

[4] C. E. Baum, J. J. Sadler, and A. P. Stone, Uniform Isotropic Dielectric Equal-Time Lenses for Matching Combinations of Plane and Spherical Waves, Sensor and Simulation Note 352, December 1992.

[5] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, AMS-55, Chapters 16 and 17, June 1964.

Reflected Laser Communication Systems

Charles M. Swenson
Assistant Professor
Utah State University

Stephen Clarke
Graduate Student
Utah State University

September 22, 1994

35-1

Reflected Laser Communication Systems

Charles M. Swenson
Assistant Professor
Utah State University

Stephen Clarke
Graduate Student
Utah State University

ABSTRACT

Under the AFOSR summer research program researchers from Utah State University / Space Dynamics Lab spent 12 weeks at the Phillips lab developing the reflected laser communication concept for low Earth orbiting satellites. The work was divided into proof of concept studies, technical demonstrations, and organization of a future cooperative research. The bench testing of a ferroelectric liquid crystal based retromodulator indicated that bandwidths on the order of 50 kHz can be achieved with off the shelf components. A proof of concept study using a high altitude balloon has been organized to further test the ferroelectric liquid crystal based retromodulator design of Utah State. This report outlines some of the reflected laser communications activities conducted under this summer research program.

Reflected Laser Communication Systems

Charles M. Swenson
Assistant Professor
Utah State University

Stephen Clarke
Graduate Student
Utah State University

1 Introduction

Satellite laser communications concepts have been under development for many years. Laser communications from ground to low earth orbit is relatively simple because most of the hardware is on the ground and the satellite need only be provided with a relatively simple sensor. However, conventional approaches to laser communication from satellite to ground require sophisticated hardware aboard the satellite such as a reliable laser subsystem and a capable subsystem for optical acquisition, tracking, and pointing to the receiver. The complexity of this hardware introduces concerns about cost, added weight, power consumption and reliability; the efforts needed to resolve these concerns have significantly delayed the implementation of laser communications systems on operational satellites. An example of the effects of this problem is the recent cancellation of the DSP Laser cross link program after 10 years of development and approximately $600 million spent.

An alternate approach to laser communications which does not require extensive hardware on the satellite is possible for satellite down links. As

stated above, the problem of most concern is communications from satellite to Earth. The alternate concept is as follows: First a laser site on the ground illuminates the satellite with either a continuous beam or a uniform train of pulses. On board the satellite is a corner cube, which reflects the incident laser light back to the transmitter on the ground. Associated with the corner cube is a modulation mechanism which reacts rapidly to change the intensity or the polarization of individual reflected pulses. By appropriately controlling this retromodulator the laser light pulses reflected back to the ground station are encoded with the communications signal (see Figure 1).

This approach offers the potential advantage of significantly reducing the size, power, and weight of hardware required aboard the satellite for laser communications while maintaining the advantages of laser communications in general. Reflected laser communications is a relatively new concept that has many potential applications. This technology may be appropriate where:

- The remote part of the communications link must be *very low power.*

- The remote site is inaccessible except by free space transmission.

- Weight and volume at the remote sight are a concern.

- There is a need for a "secure" data link.

- RF jamming is a concern.

- Simplicity at the remote is desirable.

- A moderate data rate is required.

- Occasional delayed access to communications are acceptable.

- There is a need for simultaneous communications access to the remote sight.

A few specific applications that could benefit from such a technique are:

- Communications during atmospheric re-entry.

- Communications to/from downed pilots.

- Satellite to satellite links.

- Satellite to earth links:

    - Redundant or backup communications link.
    - Fly along noninterfering link.
    - Primary communications link.

This type of communications system is very attractive for small or micro satellites since overall power concerns are primary in these simple satellite designs.

Utah State University / Space Dynamics Lab is currently involved in reflected laser communications research and has produced feasibility studies, link budgets, and tested key parts of a retromodulator design [*Jensen and Swenson*, 1992]. Dr. Charles Swenson and Stephen Clarke, graduate student, from the department of electrical engineering at Utah State University have spent approximately 12 weeks during the summer of 1994 working on proof of concept studies and technical demonstrations of the reflected laser communication concept at Phillips Lab under the Air Force Office of Scientific Research Summer Research Program. The work was concentrated in two directions. First, a bench test of key components of a retromodulator

and second, organization of a balloon based demonstration of a reflected laser communication system. A proposal to NASA to test the reflected laser communication as a cooperative effort between NASA, USU, and the Phillips lab was also generated and submitted this summer. This report outlines some of the reflected laser communications activities conducted under this summer research program.

2   Retromodulator development

The key to an effective reflected laser communications link is the retromodulator and this is where most of the initial development effort has been directed. A retromodulator for long distance communications (eg. low Earth orbit to ground) needs to be a rather large aperture device (more than 1 inch diameter), with a large acceptance angle. In addition, this device must be able to modulate at high speeds. For use on a spacecraft, the modulator must be a low power device. Currently, the most promising off the self device is a ferroelectric liquid crystal (FLC) which is very similar to the common liquid crystal technology used in consumer electronics. This family of devices has been space qualified and we foresee no problems space qualifying FLCs which are extremely low power devices. The planned balloon experiment will be testing retromodulators built from FLCs. Communications links up to 100 kbps appear possible with these devices which can be made up to 3 inches in diameter and have acceptance angles as large as 45 degrees. There is evidence that with continued development FLC's could be developed for used in links up to 1 Mbps. We note that there are other technologies in development at Sandia National Labs for retromodulators
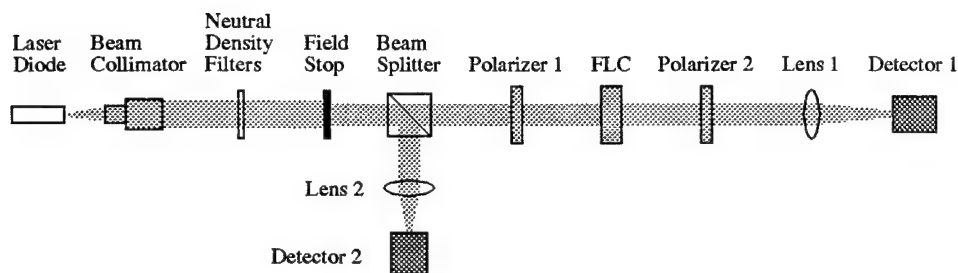
Figure 1: Experimental layout for testing FLC light valve operations.

that would make 10 to 100 Mbps links possible.

2.1 Testing

Six one inch ferroelectric liquid crystal half wave plates (FLC) were purchased from Displaytech Inc. for laboratory bench testing. Unfortunately there was a 9 week paperwork delay in obtaining the half wave plates and only a limited amount of testing was completed in the remaining time. The FLCs were placed between two crystal polarizers and illuminated with a collimated 810 nm laser diode. The experimental layout is illustrated in Fig. 1.

Some the issues examined and the observations from the testing were:

- Speed: The FLCs show usable modulation depth out to 50 kHz. A graph of the normalized transmission verses modulation rate for several FLCs is shown in Fig. 2. We note that the average transmission for each FLC varies and can be controlled by polarizer orientation and bias. The effect and its control particularly need further study.

- Temperature effects: It was found that neither the speed nor the modulation depth was greatly effected within a 20° to 40° C temperature
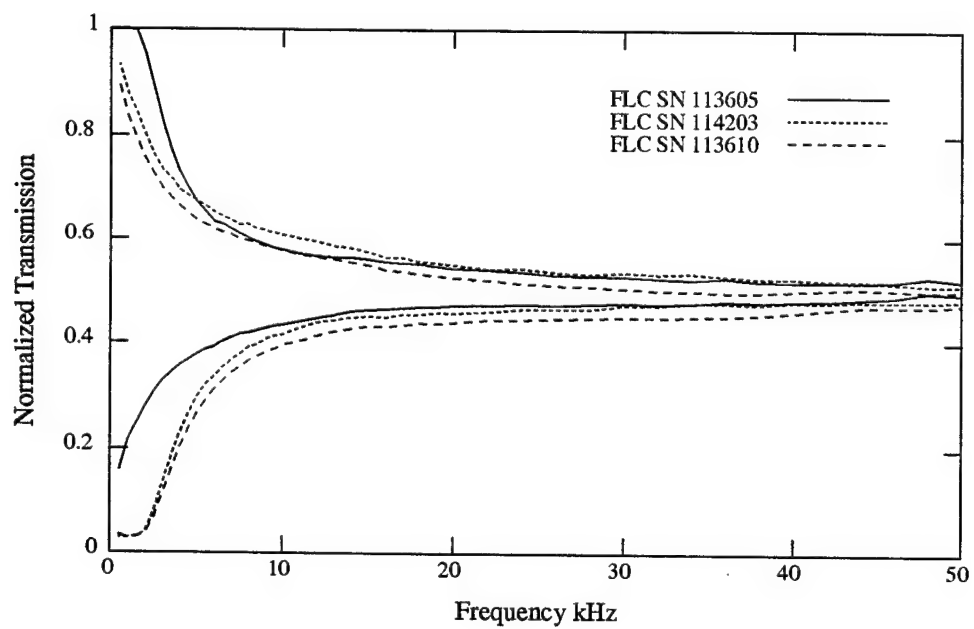
Figure 2: Transmission verses modulation speed for several different FLCs.

range.

- Biasing effects: A slight DC bias of less than 1 Volt effects the modulation depth and average transmission at high speeds.

- Polarizer orientation: the orientation of the polarizers greatly effects the modulation depth and the average transmission at high speeds. We found that maximum modulation depth at high speed did not occur at with both polarizers aligned as expected, but occurred at an angle ranging between 30 to 80 degrees. This orientation correlated with the average transmission being the mean of the max and min modulation depth at low speed.

- Defect Effects: The FLCs we tested had various aperture defects that made them unsuitable for low speed high contrast applications typical for FLCs. None of these defects were a detriment to the retromodulator application.

The final bench testing done with the FLC's was to demonstrate them in the retromodulator configuration. The setup for this experiment is illustrated in Fig. 3. The retromodulator concept worked as expected but unfortunately there was little time remaining for testing. The experimental setup was only qualitatively tested to determine if the modulated return signal could be detected and if it was significantly different than expected. The FLC retromodulator configuration generally operated as expected. Detailed testing is needed.
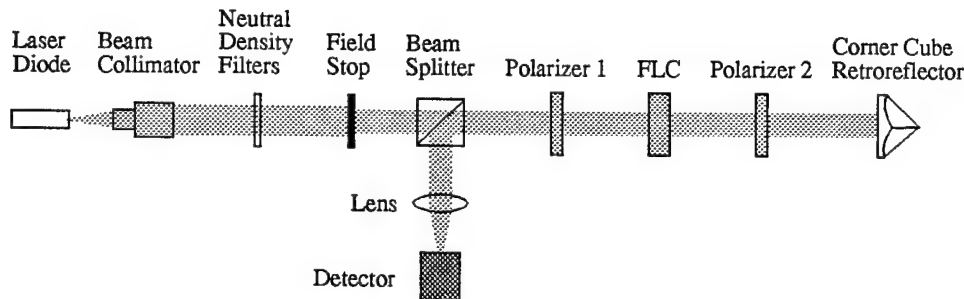
Figure 3: Experimental layout for testing FLC light valve operations.

3 Balloon Test Program

PL/SX is launching a balloon payload in January of 1995 for Sandia National Labs. The primary mission is atmospheric sampling but there is room for an additional payload providing it is small and can be easily mounted on the existing package. USU and Phillips lab have recently teamed on a proposal to USRA for the STEDI small satellite program to include a test of the reflected laser communications concept. USU would provide the spacecraft portion of the hardware and Phillips Lab would provide the needed ground station capabilities. Jeff West of PL/SX suggested that this balloon flight could be used to provide a simple proof of concept test of the proposed satellite project. A balloon at 24 km altitude is probably the next closest platform to an actual satellite as most of the turbulent effects of the atmosphere would be observed. The appears to be an ideal opportunity for a quick and relatively inexpensive demonstration of the reflected laser communications concept as well as a good "dry run" for a potential satellite demonstration.

3.1 Program Objectives

Here we present a brief outline of the objectives of the balloon test.

1. Develop a working relationship between USU and PL on reflected laser communications.

2. Demonstrate the basic principles of the reflected laser technique in a way that show scaleability to a satellite.

3. Test atmospheric effects (scintillation, backscatter, etc) on vertical propagation through the atmosphere.

4. Validate the communications link budget for reflected communications.

5. Test an assembly of FLC based retromodulators for parallel operation, temperature effects, speed, and modulation depth.

3.2 Program Description

The balloon will be launched from Kirtland Air force Base's auxiliary field heliport. Ideally the balloon will be launched 3 to 4 hours before sunrise and the experiment will be carried out both in darkness and during sunrise. The balloon will drift east for a period of about 2.5 hours until it reaches an altitude of 85,000 ft at which point the high altitude winds will bring the balloon back westward. Approximately 4 hours after the balloon launch the payload will be within a 60 degree viewing angle from vertical for SOR. Observations will continue for the next two hours as the viewing angle decreases to approximately 15 degrees. During this period SOR will illuminate the payload with a low-power laser and look for modulated returns.

The retroreflector assembly will be similar to Fig. 4 and will provide at least a 60 degree acceptance angle or a 120 degree full angle field of view. The retroreflector payload will cycle through a test pattern approximately every minute. This test pattern will consist of a bit sequence transmitted multiple times at several rates. At the lowest rate the FLC will have a modulation depth of 1 whereas at the highest rate of 100 kHz the modulation depth will approach 0. The objective is to capture two or three of these test patterns for analysis.

Utah State University will be responsible for the over all experiment design and will provide the electronics for generating the test pattern, driving the FLCs, and controlling their temperature. The electronics package will be built in the labs of PL/VT with cooperation and support of VT. Temperature data will be stored throughout the flight in solid state memory for later analysis requiring recovery of an undamaged payload for data retrieval. The analysis of the onboard and SOR data will be done at Utah State University in with cooperation with SOR. The final report will be prepared by Dr. Charles Swenson.

PL/SX will provide the balloon launch, design and fabricate the mechanical aspects of the retromodulator experiment, and provide a mechanism for attaching the payload to the balloon. The design of the retromodulator housing will be in cooperation with Utah State University.

PL/LI will provide the retroreflector and the FLC's. These will be be installed in the retromodulator housing and integrated with the electronics package by Utah State University. LI will provide the ground station resources through SOR for this experiment.

TOP VIEW

RETROMODULATOR
WINDOW

CONNECTOR
TO ELECTRONICS
BOX

RECEIVER

0.5"

5.0"

2.0"

5.0"

SIDE VIEW

MODULATOR / POLARIZERS

FILTER

RETROREFLECTOR

2.25"
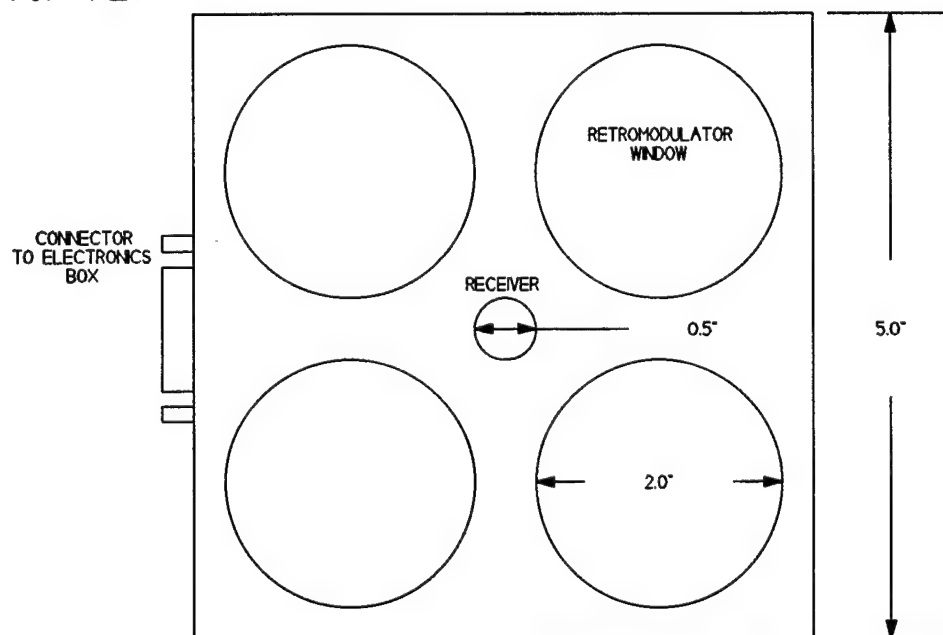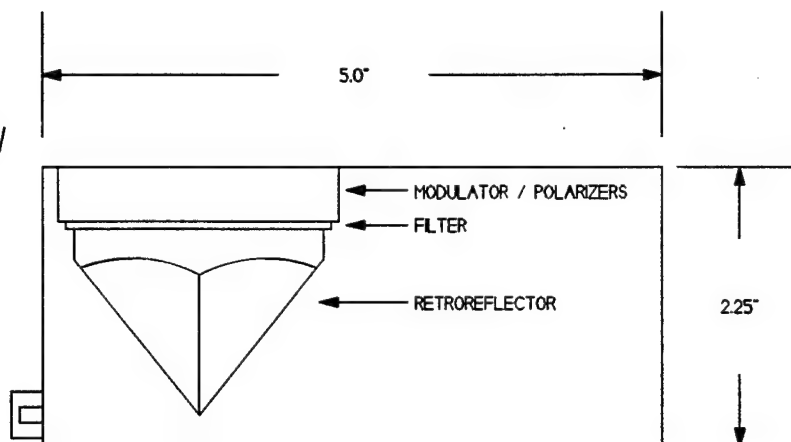
Figure 4: Experimental retromodulator assembly for a satellite test of the FLC retromodulators.

3.3 Balloon Electronics

The goal of this portion of the project was to design and build an electronics support package for the projected balloon launch. The requirements of the electronics package are summarized into the following four areas.

- A variable modulation rate, from 1kHz to 100kHz.

  This was accomplished using the built-in programmable UART on the Dallas Semiconductor 80C320 microcontroller. The baud rate is programmed to begin at 1k bits per second and end at 100k bits per second. From 1k bits per second to 10k bits per second the step size is by 1k bits. From 10k bits per second to 50k bits per second the step size is by 5k bits. From 50k bits per second to 100k bits per second a single modulation occurs at 75k bits per second. This gives a total of 20 different modulation rates.

- Test pattern for the retromodulators:

  The bit pattern selected consists of three basic patterns. First an alternating set of logic ones and logic zeros. Second a string of logic zeros with a single logic one in the center. Third is the opposite of the second, a string of logic ones with a single logic zero in the center. Each test pattern consists of sixteen bits. The test pattern is repeated twice at each of the modulation rates.

- Temperature control of the electronics and the FLCs:

  The FLCs operating performance is directly related to their temperature. Therefore the temperature of the individual FLCs will be monitored. If the temperature of an FLC falls to a critical level the heating

element for that FLC will be turned on, until the FLC returns to a good operating temperature.

- Recording the temperature of the FLCs for post flight analysis:

  Since the FLC temperature is critical a record of the temperature of each FLC will be kept for the duration of the balloon flight. The temperature profile will be stored in a non-volatile memory, as well as a time stamp of when the temperatures determined. This data will be downloaded to a computer for analysis after the flight.

At the core of the electronics package is a Dallas Semiconductor DS80C320 8-bit microcontroller. A schematic drawing of the proposed balloon experiment electronics is shown in Fig. 5. The DS80C320 is instruction set compatible with the Intel 8051. Some of the main features that led to the selection of this device are the dual internal full-duplex UARTs, multiple 16-bit timers, and an optimized architecture that reduces the average time per instruction. One of the UARTs is used to the modulate the FLCs, and the other is used as a serial communications link between the electronics package and a computer.

The temperature sensors used are the AD590. The AD590 is a two terminal integrated circuit temperature transducer which produces an output current proportional to absolute temperature. For supply voltages between +4V and +30V the device acts as a high impedance, constant current regulator passing 1uA/K. The AD590s were selected because of their good linearity over a large temperature range. The output current from the AD590 is fed through a 1k ohm, 1% tolerance resistor. The voltage developed across this resistor is directly proportional to the temperature. A single 8-bit, 16 input
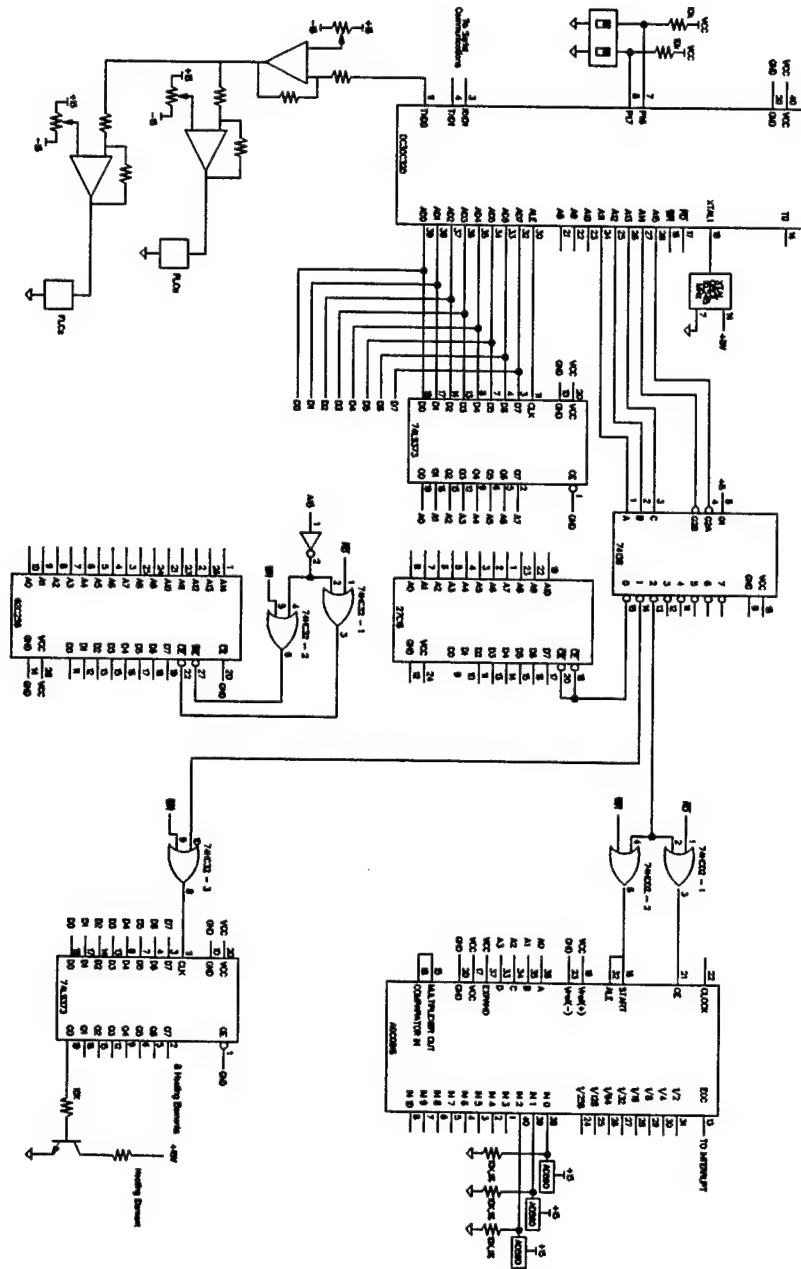
Figure 5: Schematic diagram of the electronics for the planned balloon experiment.

analog-to-digital converter is used to convert the voltage across the resistor in to its digital representation.

The temperature data is stored in a 32k x 8 SRAM. This is more than adequate storage for the expected six hour balloon flight. The SRAM sits in a socket, the DS1213D from Dallas Semiconductor, that contains an embedded lithium battery that is used to retain the data. Power-fail detection occurs between 4.75 and 4.5 volts. When Vcc falls below 4.75 volts, the chip enable is inhibited. This prevents false data from being written to the SRAM. Also the internal battery supply is switched on, holding the current contents of the SRAM. The embedded lithium battery has a data retention time of greater than 10 years.

The heating elements are controlled through an 8-bit latch. This allows one, all, or any combination of the heaters to be active at a time. The current for each heating element is provided by an external transistor connected to the latch. This allows the current through the heater to be maximized while keeping the latches within operating tolerances.

The signal coming out of the UART is at TTL levels and needs to be amplified to meet the operating requirements of the FLCs. Therefore the output of the UART is connected to an op-amp with a gain of 4, and a -10 volt offset. This gives a voltage swing from -10 to 10 volts. The output of this op-amp if fed into a series of op-amps, one for each FLC. This second op-amp allows the voltage swing and the offset to be optimized to each FLC. From bench testing the FLCs appear to work best with a -10 to 10 volt signal and a 0.4 volt offset, this offset varies slightly with modulating frequency.

A set of dip switches is used to select the mode of operation of the electronics package. There are three modes of operation, first one to clear

the memory and set the time on the time keeping socket. A second to transmit the data stored in the memory through the UART to a computer. And third the modulation and acquisition mode. This design is currently being bread-boarded and tested prior to layout of a printed circuit board.

4 Summary

The twelve weeks of the AFOSR summer research program have been very productive in developing a cooperative relationship between three directorates (VT, SX, LI) within Phillips lab. Significant progress has been made in arranging for the Starfire optical range to support future reflected laser communications projects. A significant proof of concept study has been arranged between USU and Phillips lab for this coming January by making use of extra payload capability of a high altitude balloon. Significant bench test results were obtained on FLC based retromodulators that can be continued at Utah State University. The AFOSR summer research program is a very important and effective program in promoting cooperative research between Universities and the Air Force labs.

References

Jensen, G., and C. Swenson, 1992. A laser downlink for small satellites using an optically modulated retroreflector. In *Proc. 6th Annual AIAA/USU Conf Small Sat.*

A MATHEMATICAL MODEL OF
SELF SIMILAR COMPRESSION OF COMPACT TOROIDS

Y. Chia Thio
Associate Professor
Department of Physics

with
R. Peterkin[*], Tao Du[**], and G. Kiuttu[*]

University of Miami
P. O. Box 248046
Coral Gables, FL 33124-0530

(305) 284-3150
e-mail: thio@phyvax.ir.miami.edu

[*] Phillips Laboratory, Kirtland AFB, NM.
[**] Physics Department, University of Miami, Coral Gables, FL.

A MATHEMATICAL MODEL OF SELF COMPRESSION OF COMPACT TOROIDS

Y. Chia Thio
Associate Professor
Physics Department
University of Miami, Coral Gables, Florida

Abstract

A mathematical model using a lumped parameter approach is developed in a systematic and rigorous manner to study the self similar compression of compact toroids (CT) in a coaxial and conical railgun.The CT is modelled as a ring with annular spherical caps as its bottom and top boundaries orthogonal to the radius vector rather than planar boundaries orthogonal to the z-axis. The ring element is endowed with the appropriate internal plasma flow with both radial and azimuthal kinetic energy and momentum. It is further endowed with an appropriate internal energy and a magnetic field. The presence of an initial magnetic field, an internal energy and rotational kinetic energy have immense effects on the compression dynamics of the CT. These effects are rigorously modelled and studied with respect to the stagnation points at which the maximum radial convergence of the CT is obtained under a particular set of operating conditions. We found that the effective inductance gradient of the accelerator, L', is sensitive to the precise shape of the drive current. In the present paper, the drive current is assumed to flow in the q-direction along the upper spherical cap.

The development model has been very rewarding in several ways. It leads directly to closed form analytical expressions describing the physics of the CT compression explicitly in several major cases. The closed form analytical solutions are found to be very useful for studying quantitatively the details of the compression dynamics. The mathematical development leads naturally to simple computer models for simulating the operation at the system level of actual devices using realistic power supplies including capacitor banks, inductor with opening switch, and magnetic flux compression generators. This should provide a very useful aid to the engineering design of actual CT compression devices and to the interpretation of experiments.

A MATHEMATICAL MODEL OF SELF SIMILAR COMPRESSION OF COMPACT TOROIDS

Y. Chia Thio, R. Peterkin, Tao Du, G. Kiuttu

1 Background

Current experimental results have led us to explore the compression of compact toroids (CT) in a self similar manner in a conical coaxial railgun (Figure 1.1) [1],[2],[3]. The practical significance of self similar compression lies in the fact that this mode of compression is geometrically the most efficient. That is, complete three dimensional compression of the toroid is achieved, resulting in the plasma density increasing inversely with the third power of its radius.
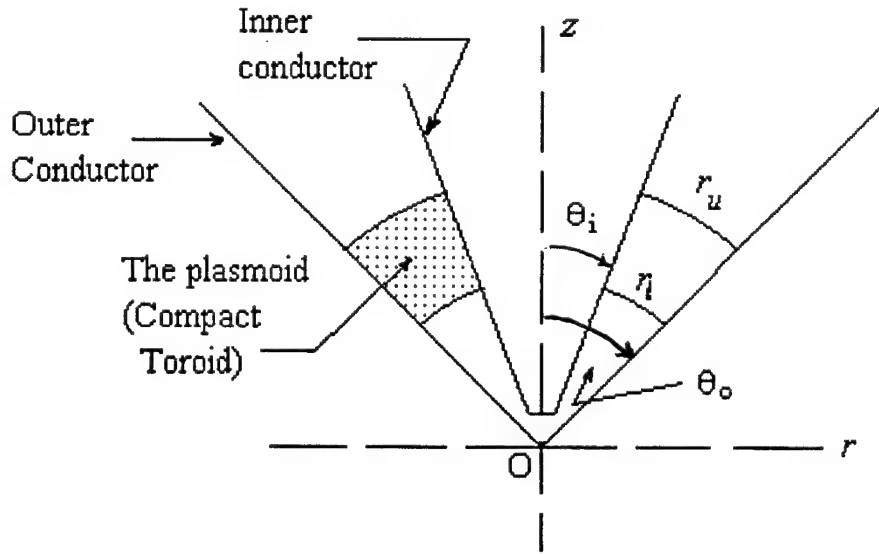
Figure 1.1: Compression of plasmoids using a conical coaxial railgun

Peterkin[3] conducted a series of full $2\frac{1}{2}$-D MHD (MACH2) simulations of the compression dynamics to investigate the physical conditions under which the compact toroids can be compressed in a self-similar fashion. He found empirically from his numerical experiments that, the compression dynamics admits a similarity variable in the form of a dimensionless parameter,

$$P = \left| \frac{mr\,\ddot{r}}{E_B} \right| \tag{1.1}$$

where m is the mass of the CT, r is its radius, $\ddot{r}$ is its instantaneous acceleration, and E_B is the total magnetic energy embedded in the CT. He found that, for values of P in the range of 1-10, self-similar compression of the CT is possible.

In this paper, our aim is to develop a simple, lumped element, model of the self-similar compression dynamics as an aid in the design and the interpretation of the compression experiments. The model yields to closed-form solutions of the equation of motion of the CT in special cases. These analytical results are used to study quantitatively the dynamics of the compression of the CT.

The basic scheme is illustrated in Figure 1.1. The CT arrives at the entrance to the conical compressor with some initial velocity, magnetic flux and internal energy. An external current flows axially along the inner conductor, across a thin layer of plasma at the rear of the CT, and return via the outer conductor.

The current carrying plasma acts as an armature, a plasma armature, like in an electric motor. The current in the conductors generates a high magnetic field behind the plasma armature current. The interaction of this magnetic field with the armature current produces a $\mathbf{J} \times \mathbf{B}$ force in the armature. The plasma armature in turn pushes on the CT as the payload, accelerating and compressing it.

Our plan for this paper is to develop first the mathematics governing the compression of the plasmoid in which an initial magnetic field is embedded. In Section 2 we develop the results relating to the geometry and kinematics of the compression. In Section 3 we develop the essential physical attributes of the compact toroid as a lumped element. In Section 4, the energetics of the compression process leading to the equation of motion is presented. The integration of the equation of motion in special cases for specific forms of current drives leading to closed-form analytical solutions are presented in Section 5. Typical graphical results are presented in Section 6. In Section 7, extension of the model to the case where the compression is driven by realistic power supplies are presented, including capacitor banks, inductors with opening switch and magnetic flux compression generators.

2 Geometry and Kinematics

We shall use spherical polars unless otherwise indicated.

From experiments[5] and $2\frac{1}{2}$-D numerical simulations, we learned that the CT in the compression cone tends to form a ring with spherical caps as its upper and lower boundaries, and that the driving current flows approximately in the θ-direction along these boundaries. These features are reflected in Figure 1.1.

We investigate the dynamics of the plasmoid assuming the compression is self similar. That is, the plasmoid maintains the same shape throughout the process.

With reference to Figure 1.1, θ_i is the inner angle of the compression cone, θ_o is the outer angle, r_ℓ is the spherical radius of the lower boundary of the plasmoid, r_u is the radius of its upper boundary. Given the conical angles, the shape of the plasmoid is completely determined by the ratio,

$$s = \frac{r_u}{r_\ell} \tag{2.1}$$

Mathematically, the self similar compression of the plasmoid is characterized by a constant value of s throughout the process. The volume of the plasmoid may be calculated as:

$$\begin{aligned} V &= 2\pi \int_{\theta_i}^{\theta_o} d\theta \sin\theta \int_{r_\ell}^{r_*} dr\, r^2 \\ &= \frac{2\pi}{3} \left(s^3 - 1\right) \left(\cos\theta_i - \cos\theta_o\right) r_\ell^3 \end{aligned} \tag{2.2}$$

In this expression, only r_ℓ varies with time. The mean density of the plasmoid is, accordingly,

$$\rho = \frac{m}{\frac{2\pi}{3} \left(s^3 - 1\right) \left(\cos\theta_i - \cos\theta_o\right) r_\ell^3} \tag{2.3}$$

where m is the total mass of the plasmoid. Differentiating with respect to time, t, we obtain,

$$\frac{\dot{\rho}}{\rho} = -\frac{3\,\dot{r_\ell}}{r_\ell} \tag{2.4}$$

If we assume that the density of the plasmoid is more or less uniform, then ρ is the local density of the plasmoid at any point internal to the plasmoid. The internal flow of the plasmoid must satisfy the mass continuity equation:

$$\frac{d\rho}{dt} + \rho\, div\,\mathbf{v} = 0 \tag{2.5}$$

or,

$$div\,\mathbf{v} = -\frac{\dot{\rho}}{\rho} \tag{2.6}$$

Using (2.4) in (2.6), it follows that,

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 v_r\right) + \frac{1}{r\sin\theta} \frac{\partial}{\partial\theta} \left(v_\theta \sin\theta\right) + \frac{1}{r\sin\theta} \frac{\partial v_\varphi}{\partial\varphi} = \frac{3\,\dot{r_\ell}}{r_\ell} \tag{2.7}$$

Even with this constraint, there exists an infinity of flow patterns which are consistent with (2.7), i.e. there exist an infinity of flow patterns which is consistent with our assumption of self-similarity for the compression of the plasmoid. For simplicity, we idealize the flow such that,

$$v_\theta = 0 \tag{2.8}$$

We also assume the flow pattern to be axisymmetric, then

$$\frac{\partial}{\partial \varphi} = 0 \tag{2.9}$$

With these assumptions, (2.7) may be solved to give,

$$v_r = \alpha r + \frac{C}{r^2}$$

where,

$$\alpha = \frac{\dot{r_\ell}}{r_\ell} \tag{2.10}$$

and C is a constant of integration. To avoid the singularity at $r = 0$, we must have $C = 0$. Thus, the flow field within the plasmoid is idealized as,

$$v_r = \alpha r \tag{2.11}$$

During the formation of the toroid, the movement of the plasma through the field lines generate a rotational motion in the toroid. This rotational motion may become a significant part of the compression dynamics under high degree of compression. We allow for the effect of this rotation in our model. By the assumption of axisymmetry, (2.9), the azimuthal flow velocity,

$$v_\varphi = v_\varphi(r, \theta)$$

Assuming that the field lines are frozen in with the fluid, the field lines co-rotates with the fluid. For this to occur without destroying the Wotjer-Wells-Taylor's state of the CT, the CT must rotates like a rigid body, that is,

$$v_\varphi = \omega r \sin \theta \tag{2.12}$$

where ω is a constant for a given position of the CT. We have chosen to characterize the position of the CT by the r-coordinate of its lower boundary (r_ℓ), so we may express the azimuthal velocity as,

$$v_\varphi(r, r_\ell) = \omega(r_\ell)\, r \sin \theta$$

3 The Compact Toroid as a Lumped Element

The plasmoid may be treated as a lumped element with a translational and a rotational momentum, a kinetic energy, a thermal internal energy and a magnetic energy. The total translational momentum of the plasmoid in the **axial direction** may be computed as:

$$
\begin{aligned}
P_z &= \int v_z\, \rho dV \\
&= \tfrac{\pi}{4}\rho\left(s^4 - 1\right)\alpha r_\ell^4 \left(\cos^2\theta_i - \cos^2\theta_o\right) \\
&= \tfrac{\pi}{4}\left(s^4 - 1\right)\left(\cos^2\theta_i - \cos^2\theta_o\right)\rho r_\ell^3\, \dot{r_\ell}
\end{aligned} \tag{3.1}
$$

giving the total translational momemtum in the form,

$$\mathbf{P} = P_z\, \mathbf{e}_z, \qquad P_z = m_p\, \dot{r_\ell}$$

where,

$$m_p = \frac{\pi}{4}\left(s^4 - 1\right)\left(\cos^2\theta_i - \cos^2\theta_o\right)\rho r_\ell^3 \tag{3.2}$$

may be thought of as an equivalent translational mass of the CT. Since $\rho r_\ell^3 \propto$ mass of the plasmoid which is assumed constant, m_p is a constant of motion and is proportional to the mass of the plasmoid.

The angular momentum of the plasmoid about the z-axis is,

$$
\begin{aligned}
I_\omega &= \int dm \, v_\varphi \, r \sin\theta \\
&= 2\pi\rho\omega \int_{\theta_i}^{\theta_o} d\theta \, \sin^3\theta \int_{r_\ell}^{r_*} dr \, r^4 \\
&= \frac{2\pi}{5} \left(s^5 - 1\right) \left[\cos\theta - \frac{\cos^3\theta}{3}\right]_{\theta_o}^{\theta_i} \rho r_\ell^3 \omega r_\ell^2
\end{aligned}
$$

giving the angular momentum in the form,

$$
I_\omega = m_\omega r_\ell^2 \omega \tag{3.3}
$$

where,

$$
m_\omega = \frac{2\pi}{5} \left(s^5 - 1\right) \left[\cos\theta - \frac{\cos^3\theta}{3}\right]_{\theta_o}^{\theta_i} \rho r_\ell^3 \tag{3.4}
$$

may be considered as an equivalent rotational mass and is a constant since it is proportional to the mass of the plasmoid. Moreover, neglecting any azimuthal skin drag, there is no net external torque on the plasmoid. Thus, the total angular momentum I_ω of the plasmoid is constant. From (3.3), we see that the plasmoid spins up as it is compressed with its angular frequency goes up inversely as the square of the radial position of the plasmoid,

$$
\omega = \left(\frac{I_\omega}{m_\omega}\right) \frac{1}{r_\ell^2}
$$

The total kinetic energy of the plasmoid may be computed in a similar way as:

$$
\begin{aligned}
E_k &= \int \tfrac{1}{2}\rho \left(v_r^2 + v_\varphi^2\right) dV, \quad \mathbf{v} = v_r \mathbf{e}_r + v_\varphi \mathbf{e}_\varphi \\
&= \tfrac{1}{2}\rho \iiint \left(\alpha^2 r^2 + \omega^2 r^2 \sin^2\theta\right) \, r^2 \sin\theta \, d\theta \, d\phi \, dr \\
&= E_r + E_\varphi
\end{aligned} \tag{3.5}
$$

where E_r is the part of the kinetic energy associated with the radial (translational) velocity, and E_φ the azimuthal velocity. The radial part may be computed as,

$$
\begin{aligned}
E_r &= \tfrac{1}{2}\rho \iiint \left(\alpha^2 r^2\right) \, r^2 \sin\theta \, d\theta \, d\phi \, dr \\
&= \frac{\pi}{5} \left(s^5 - 1\right) \left[\cos\theta_i - \cos\theta_o\right] \rho r_\ell^3 \, \dot{r}_\ell^2
\end{aligned} \tag{3.6}
$$

giving E_r in the form,

$$
E_r = \frac{1}{2} m_r \, \dot{r}_\ell^2
$$

where,

$$
m_r = \frac{2\pi}{5} \left(s^5 - 1\right) \left[\cos\theta_i - \cos\theta_o\right] \rho r_\ell^3 \tag{3.7}
$$

may be thought of as the equivalent mass for translational kinetic energy. It is proportional to the mass of the CT and is again a constant of motion. The kinetic energy associated with the azimuthal velocity, E_φ, is given by,

$$
\begin{aligned}
E_\varphi &= \tfrac{1}{2}\rho \iiint \left(\omega^2 r^2 \sin^2\theta\right) \, r^2 \sin\theta \, d\theta \, d\phi \, dr \\
&= \frac{\pi}{5}\rho \left(s^5 - 1\right) r_\ell^5 \, \omega^2 \left[\cos\theta - \frac{\cos^3\theta}{3}\right]_{\theta_o}^{\theta_i}
\end{aligned}
$$

Comparing with (3.4), the above expression can be rewritten as,

$$
\begin{aligned}
E_\varphi &= \tfrac{1}{2} m_\omega \, r_\ell^2 \, \omega^2 \\
&= \tfrac{1}{2} \left(\frac{I_\omega^2}{m_\omega}\right) \frac{1}{r_\ell^2}
\end{aligned} \tag{3.8}
$$

Note that the rotational energy goes up as the inverse square of the radius.

The internal energy of the plasmoid is taken as,

$$E_i = \frac{\bar{n}_d}{2} kTN$$

where $\overline{n_d}$ is the effective degree of freedom of the particle species comprising the plasmoid, and N is the total number of particles in the plasmoid. Assuming the plasma to be ideal, the equation of state is,

$$p = nkT \tag{3.9}$$

Combining the above two equations, we get for the internal energy of the plasmoid,

$$E_i = \frac{\bar{n}_d m}{2} \frac{p}{\rho} \tag{3.10}$$

We assume that the compression occurs adiabatically, so

$$\frac{p}{p_0} = \left(\frac{\rho}{\rho_0}\right)^\gamma \tag{3.11}$$

(3.10) and (3.11) combine to give:

$$E_i(\rho) = E_i(\rho_0) \left(\frac{\rho}{\rho_0}\right)^{\gamma-1} \tag{3.12}$$

Since, from (2.3), the density

$$\rho \propto \frac{1}{r_\ell^3}$$

Then,

$$E_i(z_\ell) = E_i(r_{\ell,0}) \left(\frac{r_{\ell,0}}{r_\ell}\right)^{3(\gamma-1)} \tag{3.13}$$

This is the scaling law for the internal energy when the compression is adiabatic.

If the plasmoid has an embedded magnetic field, such as is the case of a compact toroid, then the plasmoid possesses an amount of magnetic energy given by:

$$E_B = \frac{B^2}{2\mu} V \tag{3.14}$$

where V is the volume of the plasmoid. Assume that the internal flux is frozen in with the internal flow of the plasmoid, then

$$\frac{B}{\rho\ell} = constant \tag{3.15}$$

where ℓ is a characteristic length of the plasmoid, such as r_ℓ. When the scaling relationship for the density is used, the above relationship yields the scaling law for the embedded magnetic field of the plasmoid to be,

$$B = B_0 \left(\frac{r_{\ell,0}}{r_\ell}\right)^2 \tag{3.16}$$

Using this in the expression for the total magnetic energy embedded yields the scaling law for the embedded magnetic energy as,

$$E_B(r_\ell) = E_B(r_{\ell,0}) \left(\frac{r_{\ell,0}}{r_\ell}\right) \tag{3.17}$$

4 The Energetics and The Equation of Motion

The overall energy balance throughout the compression process may be stated in the form,

Energy input from power source $(\triangle E_{in})$
= Gain in kinetic energy of plasmoid $(\triangle E_k)$
+ Gain in magnetic energy in the accelerator $(\triangle \frac{1}{2} L i^2)$
+ Gain in magnetic energy in the plasmoid $(\triangle E_B)$
+ Gain in internal energy of plasmoid $(\triangle E_i)$
+ Thermal conduction losses (q_{th})
+ Radiation loses (q_{rad})

That is,

$$\triangle E_{in} = \triangle E_k + \triangle(\frac{1}{2} L i^2) + \triangle E_B + \triangle E_i + q_{th} + q_{rad} \tag{4.1}$$

The energy transferred to the walls (q_{th} , q_{rad}) are important as it can induce ablation, photoelectrons, photoionization etc. However, their magnitudes are small compared to $\triangle E_B$ and $\triangle E_i$. So for the global energetics of the plasmoid, we ignore these thermal losses and approximate the energy balance equation as,

$$\triangle E_k = \triangle E_{in} - \triangle(\frac{1}{2} L i^2) - \triangle E_B - \triangle E_i \tag{4.2}$$

where L is the instantaneous inductance of the accelerator as seen by the current. Equation (4.2) is the exact energy balance equation if the compression of the plasmoid proceeds adiabatically.

The external power supply sees a back emf at the 'breech' end of the accelerator (compressor) given by,

$$e = \frac{d\varphi}{dt} = \frac{d}{dt}(Li)$$

The power input into the accelerator is thus,

$$W = i \cdot e = i \frac{d}{dt}(Li)$$

The energy input into the accelerator through the breech in a time interval $[t_0, t]$ is,

$$\triangle E_{in} = \int_{t_0}^{t} W \, dt = \int i \, d(Li)$$

$$= \int \frac{1}{2} i^2 dL + \int d\left(\frac{1}{2} L i^2\right)$$

That is,

$$\triangle E_{in} - \triangle\left(\frac{1}{2} L i^2\right) = \int \frac{1}{2} i^2 dL$$

Now, the current is assumed to flow through an infinitesimally thin plasma armature at the upper boundary of the plasmoid ($r = r_u$). Thus, an infinitesimal increment in the inductance of the current circuit, dL, due to an associated infinitesimal displacement of the plasma armature in the direction of travel of the plasmoid, $-dr_u$, is given by,

$$dL = -L' \, dr_u$$

where L' is the inductance gradient $\left(-\frac{\partial L}{\partial r}\right)$ of the accelerator. Thus,

$$\triangle E_{in} - \triangle\left(\frac{1}{2} L i^2\right) = \int -\frac{1}{2} L' i^2 \, dr_u \tag{4.3}$$
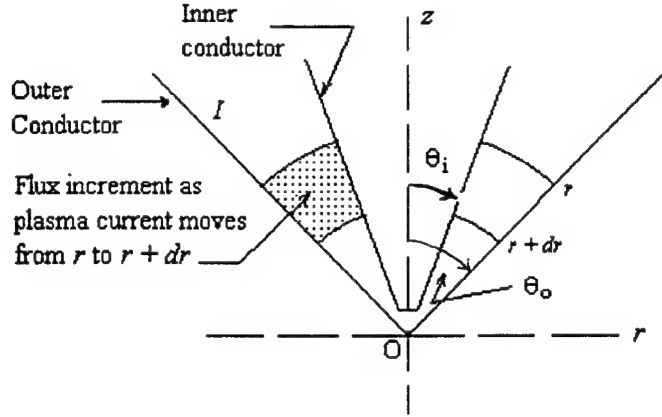
Figure 4.1: Derivation of the inductance gradient, L', of the conical plasma accelerator

The L.H.S. represents the net mechanical work done on the CT. The R.H.S. may thus be interpreted as the work integral, with $\frac{1}{2}L'i^2$ acting as the net effective force on the CT with respect to a radial displacement of the *upper boundary* of the CT.

With reference to Figure ??, the inductance gradient of the accelerator may be derived as follows. Consider a constant current flowing along the inner conductor towards the vertex of the cone (the origin) and returning via the outer conductor. Because of axisymmetry, the magnetic flux is toroidal, and the flux increment enclosed by the circuit as the armature current moves from r to $r - dr$ is,

$$
\begin{aligned}
d\varphi &= -\int_{\theta_i}^{\theta_o} B_\phi\, r\, d\theta\, dr \\
&= -r\, dr \int_{\theta_i}^{\theta_o} \frac{\mu I}{r \sin\theta}\, d\theta \\
&= -dr\, \mu I \left[\ln\left(\tan\frac{\theta}{2}\right)\right]_{\theta_i}^{\theta_o}
\end{aligned}
$$

Thus, the inductance gradient is,

$$
L' = -\frac{1}{I}\frac{\partial\varphi}{\partial r} = \mu\left[\ln\left(\tan\frac{\theta}{2}\right)\right]_{\theta_i}^{\theta_o}
\tag{4.4}
$$

Note that the expression for L' is dependent on the precise geometry of the armature current. As the shape of the armature current changes (e.g. from spherical caps to planar sheet), the value of L' changes accordingly. In this paper, we ignore any changes in the shape armature current which is assumed to remain constant as a hollow spherical cap throughout the compression process.

In summary, as the plasmoid moves from an initial position $r_{\ell,0}$ to the position r_ℓ the internal energy of the plasmoid increases by an amount,

$$
\triangle E_i = E_i(r_{\ell,0})\left[\left(\frac{r_{\ell,0}}{r_\ell}\right)^{3(\gamma-1)} - 1\right]
\tag{4.5}
$$

The embedded magnetic energy in the plasmoid increases by an amount,

$$
\triangle E_B = E_B(r_{\ell,0})\left[\left(\frac{r_{\ell,0}}{r_\ell}\right) - 1\right]
\tag{4.6}
$$

The radial kinetic energy E_r increases by an amount,

$$
\triangle E_r = \frac{1}{2}m_r\left(\dot{r_\ell}^2 - \dot{r_\ell}^2(0)\right)
\tag{4.7}
$$

36–9

The rotational kinetic energy increases by an amount,

$$\Delta E_\phi = \frac{1}{2}\left(\frac{I_\omega^2}{m_\omega}\right)\left[\frac{1}{r_\ell^2} - \frac{1}{r_{\ell,0}^2}\right] = E_\phi(0)\left[\frac{r_{\ell,0}^2}{r_\ell^2} - 1\right] \tag{4.8}$$

5 Analytical Solutions

5.1 Canonical Drive: A Special Current Profile

We shall consider an ideal current drive where the work done on the plasmoid by the external current is proportional to the increase in the magnetic energy of the plasmoid. That is,

$$i \propto E_B$$

In this case, the drive current has the functional form,

$$i = i_0 \frac{r_{\ell,0}}{r_\ell} = i_0 \frac{r_{u,0}}{r_u} \tag{5.1}$$

This current is suggested by Peterkin. Previous $2\frac{1}{2}$–D numerical simulations done by Peterkin show that this current profile produces favorable compression of CT's.

For this current drive, the work done on the CT by the external current can be computed to be given by,

$$\Delta E_{in} - \Delta\left(\frac{1}{2}Li^2\right) = \frac{1}{2}L'r_{u,0}i_0^2\left[\frac{r_{u,o}}{r_u} - 1\right] \tag{5.2}$$

This is the work done on the plasmoid. Note that this has the same scaling relationship with respect to r_ℓ as the increase in the magnetic energy of the CT (4.6),

$$\Delta E_B = E_B(z_{\ell,0})\left[\left(\frac{r_{\ell,0}}{r_\ell}\right) - 1\right]$$

Thus, with the appropriate choice of the initial current, i_0, the increase in the magnetic energy of the CT can be completely compensated by the mechanical work done on the CT. We shall refer to the choice of the external current satisfying the relationship (5.1) the case of *canonical drive.*

The above results are all that is required for evaluating the energy balance equation (4.2). For completeness, however, we shall evaluate the energy drawn from the external power supply ΔE_{in}. The increment in the magnetic energy accumulated in the accelerator as the plasma armature moves from an initial position of $r_{u,0}$ to a general position r_u can be computed as,

$$\begin{aligned}
\Delta\left(\frac{1}{2}Li^2\right) &= \frac{1}{2}L(r_u)i^2 - \frac{1}{2}L(r_{u,0})i_0^2 \\
&= \frac{1}{2}[L_0 + L'(r_{u,0} - r_u)]i_0^2\left(\frac{r_{u,o}}{r_u}\right)^2 - \frac{1}{2}L_0 i_0^2 \\
&= \frac{1}{2}(L_0 + L'_{u,0}r)i_0^2\left[\left(\frac{r_{u,o}}{r_u}\right)^2 - 1\right] \\
&\quad -\frac{1}{2}L'r_{u,0}i_0^2\left[\frac{r_{u,o}}{r_u} - 1\right]
\end{aligned}$$

where we have used the symbol L_0 to indicate the initial inductance in the accelerator, i.e. $L_0 = L(r_{u,0})$. Combining the above expression with Equation (5.2), we obtain for the energy input from the external power supply to be,

$$\Delta E_{in} = \frac{1}{2}\left(L_0 + L'r_{u,0}\right)i_0^2\left[\left(\frac{r_{u,0}}{r_u}\right)^2 - 1\right] \tag{5.3}$$

5.2 Integrating the Equation of Motion: The Special Case, $\gamma = \frac{4}{3}$, and No Rotation

Recall that, as the plasmoid moves from an initial position $r_{\ell,0}$ to the position r_ℓ, the internal energy of the plasmoid increases by an amount, (4.5),

$$\Delta E_i = E_i(r_{\ell,0})\left[\left(\frac{r_{\ell,0}}{r_\ell}\right)^{3(\gamma-1)} - 1\right]$$

If $\gamma = 4/3$, we have the happy coincidence that the internal energy increment also scales in the same way as the external work done on the plasmoid as well as its increment in magnetic energy. In this case, it is possible to fully compensate for the increase in the magnetic and internal energy of the plasmoid by choosing the external drive so that,

$$\frac{1}{2}L'r_{u,0}\,i_0^2 = E_i(r_{\ell,0}) + E_B(r_{\ell,0}) \tag{5.4}$$

Further, if there is no initial rotational energy in the plasmoid, using the compensated drive, (5.4), the energy balance equation (the first integral of the equation of motion) takes on a particular simple form,

$$\Delta E_r = 0 \tag{5.5}$$

In this case, the kinetic energy of the plasmoid remains constant. The plasmoid coasts along at constant speed. All the energy required for the compression is derived from the external power supply.

$$\dot{r}_\ell = \text{constant} = v_o, \text{ say} \tag{5.6}$$

The position of the plasmoid at time t is,

$$r_\ell = r_0 - v_o(t - t_0) \tag{5.7}$$

where we have use the symbol r_0 to indicate the initial value of r_ℓ. This notation will be used in the following unless indicated otherwise. The required pulse shape of the drive current is,

$$i = i_0\frac{r_0}{r_0 - v_o(t - t_0)} \tag{5.8}$$

5.3 Integrating the Equation of Motion: The More General Case

As the more general case, we take $\gamma = 5/3$. In that case, the internal energy of the plasmoid varies according to,

$$\Delta E_i = E_i(r_{\ell,0})\left[\left(\frac{r_{\ell,0}}{r_\ell}\right)^2 - 1\right] \tag{5.9}$$

This scales like the rotational energy.

Using all the above results, the energy equation (4.2) can be cast in the form,

$$E_r = E_T - E_d - E_B(r_{\ell,0})\left(\frac{r_{\ell,0}}{r_\ell}\right) - \{E_i(r_{\ell,0}) + E_\phi(r_{\ell,0})\}\left(\frac{r_{\ell,0}}{r_\ell}\right)^2$$

where,

$$E_T = E_r\,(r_0) + E_\phi\,(r_0) + E_B(r_0) + E_i(r_0) \tag{5.10}$$

$$E_d = \frac{1}{2}\,L'r_{u,0}\,i_0^2 \tag{5.11}$$

We introduce

$$\zeta = \frac{r_\ell}{r_{\ell,0}} \tag{5.12}$$

as the dimensionless position variable. Dividing the energy equation throughout by E_T, and recalling (3.6),

$$E_r = \frac{1}{2} m_r \, \dot{r_\ell}^2$$

the energy equation leads to an equation of motion for the CT in the form,

$$\left(\frac{m_r r_0^2}{2E_T}\right) \left(\zeta \frac{d\zeta}{dt}\right)^2 = \left(1 - \frac{E_d}{E_T}\right) \zeta^2 + \left(\frac{E_d - E_B(r_0)}{E_T}\right) \zeta - \frac{E_i(r_0) + E_\phi(r_0)}{E_T}$$

From this a characteristic time scale emerges as,

$$t_c = \sqrt{\frac{m_r r_0^2}{2E_T}} \tag{5.13}$$

Normalizing time with respect to this characteristic time scale,

$$\tau = \frac{t}{t_c} \tag{5.14}$$

and the various energies with respect to the total initial energy of the CT,

$$\varepsilon_d = \frac{E_d}{E_T} = \frac{\frac{1}{2}L' r_{u,0} \, i_0^2}{E_k(r_0) + E_B(r_0) + E_i(r_0)} \quad , \tag{5.15}$$

$$\varepsilon_B = \frac{E_B}{E_T} = \frac{E_B(r_0)}{E_k(r_0) + E_B(r_0) + E_i(r_0)} \quad , \tag{5.16}$$

$$\varepsilon_{i\phi} = \frac{E_i(r_0) + E_\phi(r_0)}{E_T} = \frac{E_i(r_0) + E_\phi(r_0)}{E_k(r_0) + E_B(r_0) + E_i(r_0)} \tag{5.17}$$

Then the equation of motion can be rewritten in the form,

$$\left(\zeta \frac{d\zeta}{d\tau}\right)^2 = (1 - \varepsilon_d) \zeta^2 + (\varepsilon_d - \varepsilon_B) \zeta - \varepsilon_{i\phi} \equiv f(\zeta) \tag{5.18}$$

Rearranging and taking square root, we obtain,

$$d\tau = -\frac{\zeta d\zeta}{\sqrt{f(\zeta)}}$$

The negative sign is chosen because the plasmoid is injected with a negative velocity along the z-axis. Integrating,

$$\tau - \tau_0 = -\int_{\zeta_0}^{\zeta} \frac{\zeta d\zeta}{\sqrt{f(\zeta)}} \tag{5.19}$$

Here $\zeta_0 = 1$.

This integral is tabulated in Gradsteyn and Ryzhik, Formula 2.222.2, pg. 73, Formula 2.261, pg. 81 and Formula 2.264, pg. 83, [4] and may be evaluated to be:

$$\tau - \tau_0 = [G(\zeta)]_\zeta^1 \tag{5.20}$$

where,

$$G(\zeta) = \begin{cases} = \frac{\sqrt{f(\zeta)}}{c} - \frac{b}{2c} \frac{1}{\sqrt{c}} \ln(2\sqrt{cf(\zeta)} + 2c\zeta + b, & c > 0 \\ = \frac{\sqrt{f(\zeta)}}{c} + \frac{b}{2c} \frac{1}{\sqrt{-c}} \arcsin \frac{2cx+b}{\sqrt{-\Delta}}, & c < 0, \Delta \equiv 4ac - b^2 < 0 \\ = \left(\frac{1}{3}f(\zeta) - a\right) \frac{2\sqrt{f(\zeta)}}{b^2}, & c = 0 \end{cases} \tag{5.21}$$

with $a = -\varepsilon_{i\phi}$, $b = \varepsilon_d - \varepsilon_B$, $c = 1 - \varepsilon_d$, so that $f(\zeta) = c\zeta^2 + b\zeta + a$.

The case $\varepsilon_d = \varepsilon_B \prec 1$ deserves some special attention. This is the case when,

$$\frac{1}{2} L' r_{u,0} i_0^2 = E_B(r_0)$$

and $c > 0$. In this case, $b = 0$, and the solution becomes, with $\tau_0 = 0$,

$$\tau = \left[\frac{\sqrt{f(\zeta)}}{c}\right]_\zeta^1 = \left[\frac{\sqrt{c\zeta^2 + a}}{c}\right]_\zeta^1$$

That is,

$$\zeta = \left(1 - 2\sqrt{c + a}\,\tau + c\tau^2\right)^{1/2}$$

or,

$$\zeta = \left(1 - 2\sqrt{1 - \varepsilon_B - \varepsilon_{i\phi}}\,\tau + (1 - \varepsilon_B)\,\tau^2\right)^{1/2}$$

That is,

$$\zeta = \left(1 - 2\sqrt{\frac{E_r(r_o)}{E_T}}\,\tau + \left(1 - \frac{E_B(r_0)}{E_T}\right)\tau^2\right)^{1/2}$$

or,

$$\zeta = \left(1 - 2\sqrt{\frac{E_r(r_o)}{E_T}}\,\tau + \left(\frac{E_k(r_o) + E_i(r_0)}{E_T}\right)\tau^2\right)^{1/2}$$

The current is then given in this case by,

$$i = i_0\left(\frac{1}{\zeta}\right) = i_0 \frac{1}{\left(1 - 2\sqrt{\frac{E_r(r_o)}{E_T}}\,\tau + \left(\frac{E_k(r_o) + E_i(r_0)}{E_T}\right)\tau^2\right)^{1/2}} \tag{5.22}$$

5.4 Stagnation Points

From (5.18), the translational kinetic energy of the CT at the position $r = r_\ell$ is given by,

$$E_r = \frac{1}{2} m_r \left(\frac{dr_\ell}{dt}\right)^2 = \frac{1}{2} \frac{m_r r_0^2}{t_c^2} \left(\frac{d\zeta}{d\tau}\right)^2 = \frac{1}{2} \frac{m_r r_0^2}{\left(\frac{m_r r_0^2}{2E_T}\right)} \left(\frac{d\zeta}{d\tau}\right)^2 = E_T \left(\frac{d\zeta}{d\tau}\right)^2 \tag{5.23}$$

The mean radial translational velocity of the CT is,

$$v = \sqrt{\frac{2E_r}{m}} = \sqrt{\frac{2E_T}{m}} \frac{d\zeta}{d\tau} \tag{5.24}$$

That is,

$$v = \sqrt{\frac{2E_T}{m}} \frac{1}{\zeta} \sqrt{(1 - \varepsilon_d)\zeta^2 + (\varepsilon_d - \varepsilon_B)\zeta - \varepsilon_{i\phi}} = \sqrt{\frac{2E_T}{m}} \frac{1}{\zeta} \sqrt{f(\zeta)}$$

The stagnation points ($v = 0$) occur at $\zeta = \zeta_s$ where ζ_s satisfies the quadratic,

$$f(\zeta) = (1 - \varepsilon_d)\zeta^2 + (\varepsilon_d - \varepsilon_B)\zeta - \varepsilon_{i\phi} = 0 \tag{5.25}$$

The roots are:

$$\zeta_s = \frac{\varepsilon_B - \varepsilon_d \pm \sqrt{(\varepsilon_B - \varepsilon_d)^2 + 4(1 - \varepsilon_d)\varepsilon_{i\phi}}}{2(1 - \varepsilon_d)} \tag{5.26}$$

The dimensionless variable ζ_s is also the inverse of the maximum radial convergence ratio achievable.
Two special cases are immediate:

(a) $\varepsilon_d = 1$. In this case, (5.25) gives the stagnation point and inverse convergence ratio as,

$$\zeta_s = \frac{\varepsilon_{i\phi}}{1 - \varepsilon_B} = \frac{E_i(r_0) + E_\phi(r_0)}{E_k(r_0) + E_i(r_0)} \tag{5.27}$$

(b) $\varepsilon_d = \varepsilon_B$. In this case, (5.25) gives the stagnation point and inverse convergence ratio as,

$$\zeta_s = \sqrt{\frac{\varepsilon_{i\phi}}{1 - \varepsilon_B}} = \sqrt{\frac{E_i(r_0) + E_\phi(r_0)}{E_k(r_0) + E_i(r_0)}} \tag{5.28}$$

Note that the two cases are related by a square root.

We may categorize the solutions into four major classes:

(1) No external drive: $i_0 = 0$. In this case, $\varepsilon_d = 0$, and consequently $c = 1$. In this case the stagnation point and the reciprocal of the maximum radial convergence ratio is,

$$\zeta_s = \frac{\varepsilon_B \pm \sqrt{\varepsilon_B^2 + 4\varepsilon_{i\phi}}}{2}$$

For positive value of ζ_s we choose the + sign for the radical. Then,

$$\begin{aligned}
\zeta_s &= \frac{\varepsilon_B + \sqrt{\varepsilon_B^2 + 4\varepsilon_{i\phi}}}{2} = \frac{\varepsilon_B}{2}\left(1 + \sqrt{1 + \frac{4\varepsilon_{i\phi}}{\varepsilon_B}}\right) \\
&\sim \varepsilon_B = \frac{E_b}{E_T} = \frac{\text{initial magnetic energy}}{\text{initial total energy}}, \quad \text{if } \varepsilon_{i\phi} \ll \varepsilon_B
\end{aligned} \tag{5.29}$$

The initial magnetic energy (the total magnetic flux) in the CT essentially determines the degree of compression achievable.

(2) Normal drive: $\varepsilon_d < 1$. The stagnation points are essentially given by (5.26). Depending upon whether $\varepsilon_d < \varepsilon_B$ or $\varepsilon_d > \varepsilon_B$, the stagnation points are given by,

$$\zeta_s = \frac{(\varepsilon_B - \varepsilon_d)}{2(1 - \varepsilon_d)}\left(1 + \sqrt{1 + \frac{4(1 - \varepsilon_d)\varepsilon_{i\phi}}{(\varepsilon_B - \varepsilon_d)^2}}\right), \quad \text{if } \varepsilon_B > \varepsilon_d$$

or,

$$\zeta_s = \frac{(\varepsilon_d - \varepsilon_B)}{2(1 - \varepsilon_d)}\left(\sqrt{1 + \frac{4(1 - \varepsilon_d)\varepsilon_{i\phi}}{(\varepsilon_d - \varepsilon_B)^2}} - 1\right), \quad \text{if } \varepsilon_B > \varepsilon_d$$

(3) Over drive: $\varepsilon_d > 1$. The expression (5.26) need to be evaluated for both signs of the radical. The larger of the two roots give the stagnation point. In this case, when $\varepsilon_B \ll 1 < \varepsilon_d$ and $\varepsilon_{i\phi} \ll 1 < \varepsilon_d$, which are often the case in practice, both roots are negative, i.e. there is no stagnation point and the CT may be driven to any degree of convergence in the model. Of course, in reality, various instabilities would occur to limit the degree of convergence achievable.

6 Analytical Results

Typical numerical results are graphed in Figures 6.1 to 6.6 below, in which graphs of the dimensionless position variable, ζ, versus the dimensionless time variable, τ, giving the time-history of the trajectory of the CT are shown. The dimensionless variable, ζ, is also the radial compression factor. The volume compression factor is thus given by ζ^3. Each of the figures presents the results for a fixed set of values of ε_B and $\varepsilon_{i\phi}$. The different curves correspond to the different values of ε_d as indicated in each figure.

7 Extension to The General Case

In practice, the form of the current is not known *a priori*. The equation of motion must be integrated together with the external circuit equations, determining the current implicitly. Recalling equations (4.2, 4.3, 4.5, 4.6, ??), the equation of motion can be written in the form,

$$\frac{1}{2}m_r \dot{r}_\ell^2 = \frac{1}{2}m_r \dot{r}_\ell^2(0) + \int -i^2 L' \, dz_u - \triangle E_B - \triangle E_i - \triangle E_\phi \tag{7.1}$$

Figure 6.1: Compression trajectory, ζ vs τ, for the case $\epsilon_B = 0.05, \epsilon_{i\phi} = 0.001$. The different curves correspond to the different values of ϵ_d as indicated.
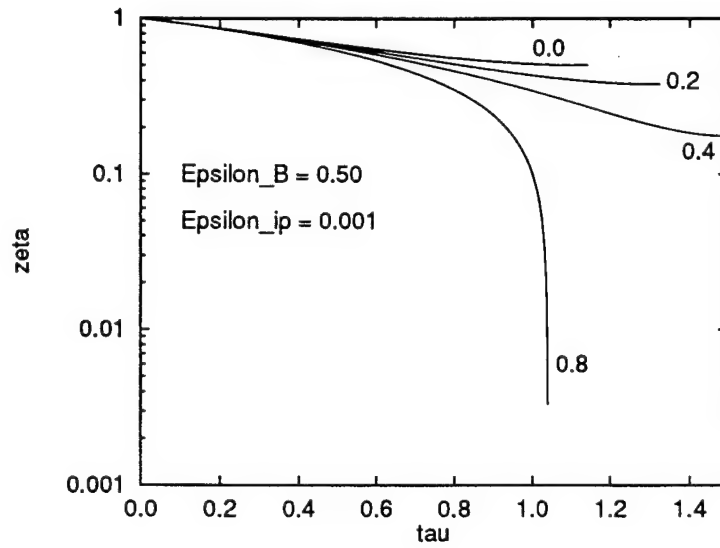
Figure 6.2: Compression trajectory, ζ vs τ, for the case $\epsilon_B = 0.5, \epsilon_{i\phi} = 0.001$. The different curves correspond to the different values of ϵ_d as indicated.
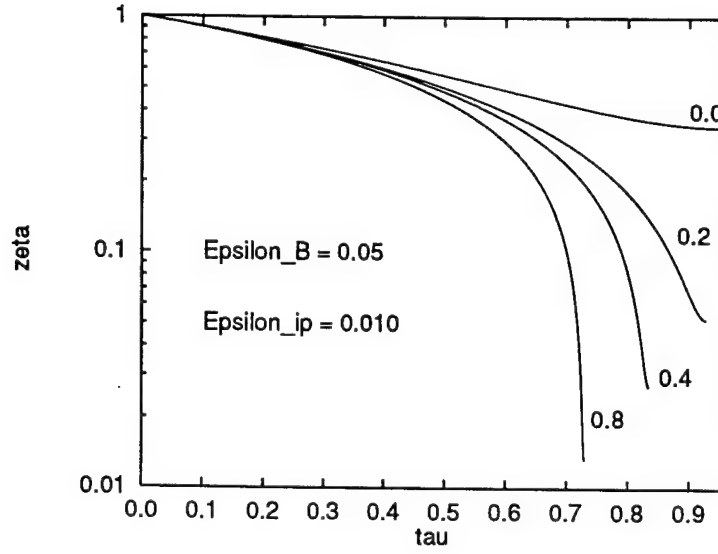
Figure 6.3: Compression trajectory, ζ vs τ, for the case $\epsilon_B = 0.05, \epsilon_{i\phi} = 0.01$. The different curves correspond to the different values of ϵ_d as indicated.
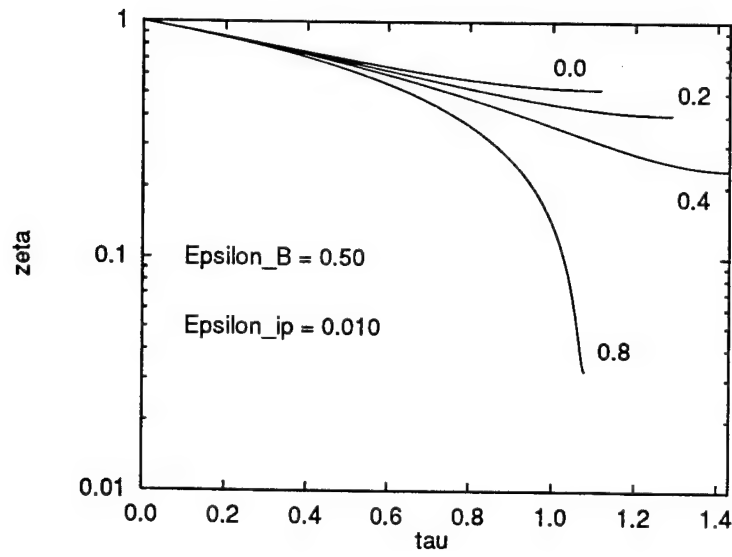
Figure 6.4: Compression trajectory, ζ vs τ, for the case $\epsilon_B = 0.5, \epsilon_{i\phi} = 0.01$. The different curves correspond to the different values of ϵ_d as indicated.
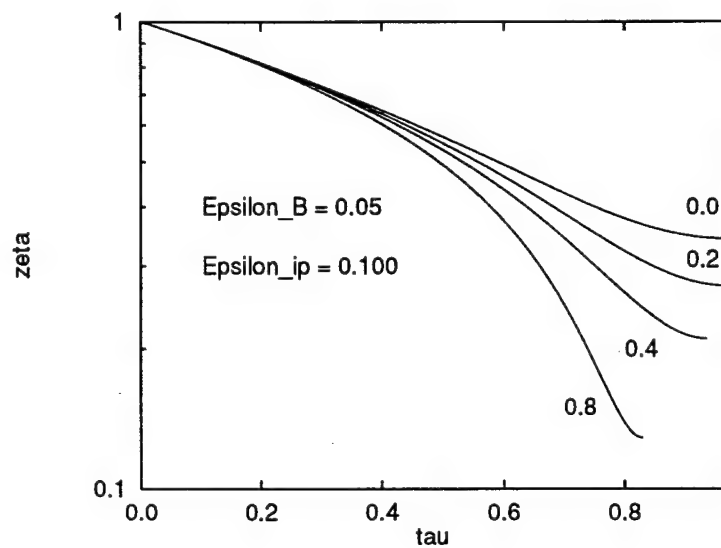
Figure 6.5: Compression trajectory, ζ vs τ, for the case $\epsilon_B = 0.05, \epsilon_{i\phi} = 0.1$. The different curves correspond to the different values of ϵ_d as indicated.
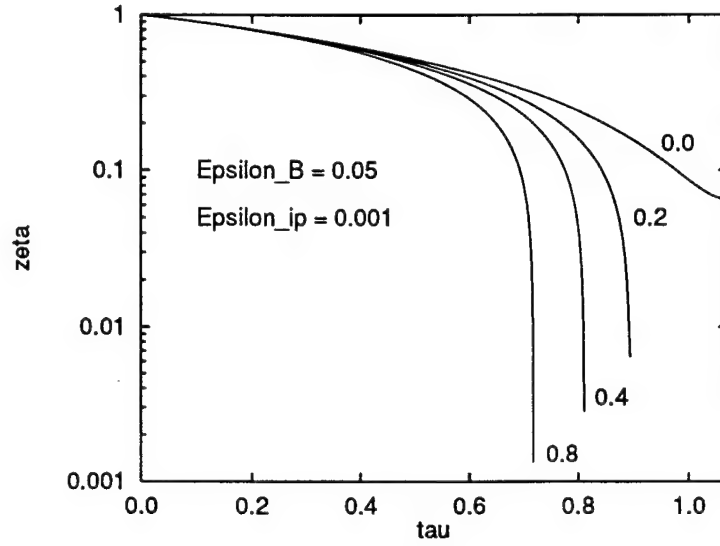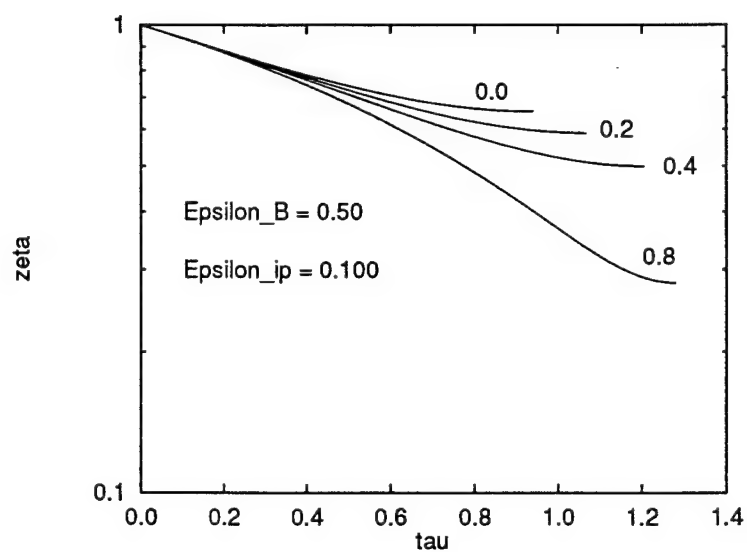

Figure 6.6: Compression trajectory, ζ vs τ, for the case $\epsilon_B = 0.5, \epsilon_{i\phi} = 0.1$. The different curves correspond to the different values of ϵ_d as indicated.

where,

$$\triangle E_B = E_B(r_0) \left[\left(\frac{r_0}{r_\ell} \right) - 1 \right]$$

$$\triangle E_i = E_i(r_0) \left[\left(\frac{r_0}{r_\ell} \right)^{3(\gamma-1)} - 1 \right]$$

$$\triangle E_\phi = E_\phi(z_0) \left[\left(\frac{r_0}{r_\ell} \right)^2 - 1 \right]$$
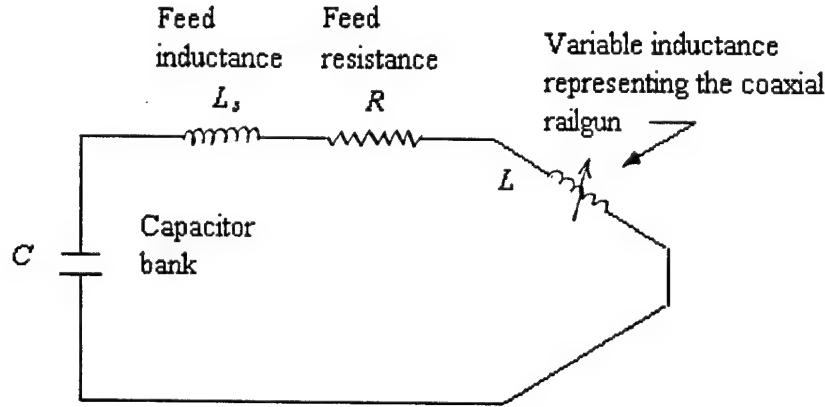
7.1 Capacitor Bank

Figure 7.1: Equivalent circuit of a capacitively driven CT compression by a coaxial railgun

If the external power supply consists of a capacitor bank with some feed inductance (L_s) and resistance (R) in series, Kirchoff's law taken around the circuit gives,

$$
\begin{aligned}
V_g &= L_s \frac{di}{dt} + Ri + \frac{d}{dt}(Li) \\
&= L_s \frac{di}{dt} + Ri - L' \frac{dr_\ell}{dt} i + L \frac{di}{dt} \\
&= L_s \frac{di}{dt} + Ri - L's \frac{dr_\ell}{dt} i + L \frac{di}{dt} \\
&= (L_s + L) \frac{di}{dt} + Ri - L's \, \dot{r}_\ell
\end{aligned}
\tag{7.2}
$$

where V_g is the instantaneous capacitor voltage. Re-arranging,

$$\frac{di}{dt} = \frac{1}{(L_s + L)} \left\{ V_g - Ri + L's \, \dot{r}_\ell \right\} \tag{7.3}$$

The capacitor voltage is related to the current by,

$$\frac{dV_g}{dt} = -\frac{i}{C} \tag{7.4}$$

where C is the capacitance of the bank. The equations (7.2), (7.3) and (7.4) form a closed set of equations which can be easily integrated by one of the standard ordinary differential equation solver (e.g. Runge-Kutta).

7.2 Inductor with Opening Switch

This is the case in which the power supply consists of a inductor, L_0, with a switch across it shunting out the load (the plasmoid accelerator) initially. An initial current is introduced into the inductor and the switch begins to open. The opening of the switch is modelled as a time-varying resistor, R_s, with
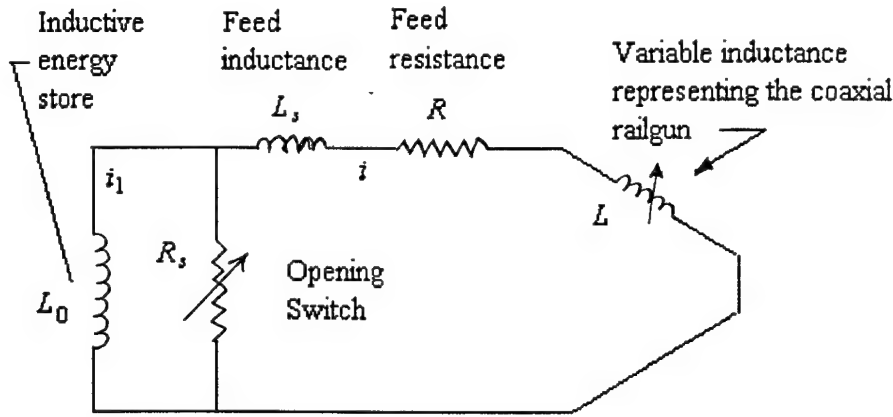
Figure 7.2: Equivalent circuit of an inductively driven CT compression by a coaxial railgun

its resistance increasing with time, thus commutating the current, i_1, in the inductor into the load. The load current, i, thus increases with time.

The circuit representation is shown in Figure 7.2. The circuit equation (7.2) applies in this case with the voltage at the terminals of the power supply V_g given by,

$$V_g = R_s(t)(i_1 - i) = -L_0 \frac{di_1}{dt} \tag{7.5}$$

The circuit equation (7.3) then becomes,

$$\frac{di}{dt} = \frac{1}{(L_s + L)} \left\{ R_s(t)(i_1 - i) - Ri + L's \, \dot{r}_\ell \right\} \tag{7.6}$$

Equations (7.5) and (7.6) may be solved together with the equation of motion (7.1) as a set of first-order o.d.e, given the switch function $R_s(t)$.

7.3 Magnetic Flux Compression Generator
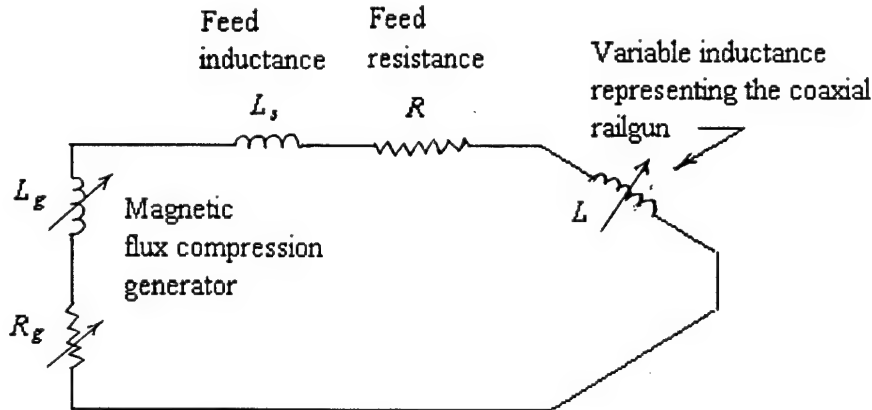
Figure 7.3: Equivalent circuit of a coaxial plasma gun driven by magnetic flux compression generator.

A magnetic flux compression generator (MCG) can be represented by a time dependent inductor in which there is an initial current. The inductance of the inductor decreases with time, squeezing out the

flux into the load, producing current in the load and transferring energy from the inductor to the load. The circuit equation (7.2) may still be applied to the circuit in which the magnetic flux compression generator replaces the capacitor bank, but now V_g is the voltage at the terminals of the generator. This voltage is generated by the inductor according to:

$$V_g = -\frac{d}{dt}(L_g i) - R_g i$$

where L_g is the instantaneous value of the inductance of the generator and R_g is its internal resistance. Thus,

$$V_g = -\left(\frac{dL_g}{dt}i + L_g\frac{di}{dt} + R_g i\right)$$

Note that for a MCG, dL_g/dt is negative and is the source of electrical energy of the generator. Substituting into equation (7.2), we get

$$\frac{dL_g}{dt}i + L_g\frac{di}{dt} + (L_s + L)\frac{di}{dt} + (R + R_g)i - L's\,\dot{r}_\ell = 0$$

$$\frac{di}{dt} = \frac{1}{(L_s + L + L_g(t))}\left\{\left(-\frac{dL_g(t)}{dt}\right)i - (R + R_g)i + L's\,\dot{r}_\ell\right\} \tag{7.7}$$

Given the generator functions $L_g(t)$ and $R_g(t)$ this circuit equation may be solved with the equation of motion (7.1).

8 Concluding Remarks

In this paper, a lumped parameter, 0-D model for the self similar compression dynamics of compact toroids have been rigorously developed. The model was found very useful in several ways. It leads directly to closed form analytical expressions describing the physics of the CT compression explicitly in several major cases. The closed form analytical solutions are found to be very useful for studying quantitatively the details of the compression dynamics. The mathematical development leads naturally to simple computer models for simulating the operation at the system level of actual devices using realistic power supplies including capacitor banks, inductor with opening switch, and magnetic flux compression generators. This should provide a very useful aid to the engineering design of actual CT compression devices and to the interpretation of experiments.

References

[1] J. H. Degnan, R. E. Peterkin, Jr., G. P. Baca, J. D. Beason, D. E. Bell, M. E. Dearborn, D. Dietz, M. R. Douglas, S. E. Englert, T. J. Englert, K. E. Hackett, J. H. Holmes, T. W. Hussey, G. F. Kiuttu, F. M. Lehr, G. J. Markin, B. W. Mullins, D. W. Price, N. F. Roderick, E. L. Ruden, C. R. Sovinec, and P. J. Turchi, "Compact toroid formation, compression, and acceleration," *Phys. Fluids B*, **5** (8), August 1993, pp. 2938-2958.

[2] R. E. Peterkin, Jr., J. H. Degnan, T. W. Hussey, N. F. Roderick, and P. J. Turchi, "A long conduction time compact torus plasma opening switch," *IEEE Trans. Plasma Sci.*, **21**, (5), Oct 1993, pp. 522-528.

[3] R. E. Peterkin, Jr., "Direct electromagnetic acceleration of a compact toroid to high density and high speed," Manuscript, April 4, 1994.

[4] I. S. Gradshteyn and I. M. Ryzhik, "Table of Integrals, Series, and Products," 4th Ed., Academic Press, New York, 1965.

[5] G. Kiuttu, private communication, July 1994.

INVESTIGATIONS OF ELECTRON INTERACTIONS WITH MOLECULES:
ELECTRON ATTACHMENT AND ELECTRON DETACHMENT REACTIONS OF HALOGENATED MOLECULES

Jane M. Van Doren
Assistant Professor
Department of Chemistry
College of the Holy Cross
1 College Street
Worcester, MA 01610-2395

and

Albert D. Kowalak
Professor
Department of Chemistry
University of Massachusetts Lowell
1 University Ave
Lowell, MA 01854

INVESTIGATIONS OF ELECTRON INTERACTIONS WITH MOLECULES: ELECTRON ATTACHMENT AND ELECTRON DETACHMENT REACTIONS OF HALOGENATED MOLECULES

Jane M. Van Doren   Albert D. Kowalak

ABSTRACT

Reactions of electrons with a variety of halogenated molecules were explored using a Flowing Afterglow Langmuir Probe and mass spectrometer. Electron attachment to octafluorotoluene, pentafluorobenzonitrile, pentafluoroacetophenone, trifluorotolunitrile, and pentafluoropyridine results in the formation of a parent anion only, i.e., the association product is formed. Rate coefficients for electron attachment to these molecules indicate that these reactions are generally efficient (at 300K, k(octafluorotoluene) = 2.1 x 10 cm^3 s^{-1}, k(pentafluorobenzonitrile) $\geq$ 2.3 x 10 cm^3 s^{-1}, k(pentafluoropyridine) = 1.8 x 10 cm^3 s^{-1}). In two cases, reactions of octafluorotoluene and pentafluoropyridine, the anion formed is observed to detach electrons under our experimental conditions. Analysis of these data assuming that detachment is thermally controlled, indicates that the electron affinities of octafluorotoluene and pentafluoropyridine are < 1 eV. Chloroacetonitrile and pentafluorobenzene react with electrons to form ions which are fragments of the parent, i.e., dissociative attachment products. Attachment to chloroacetonitrile forming Cl$^-$ is relatively efficient, k = 4 x 10 cm^3 s^{-1} at 300K and becomes increasingly more efficient as temperature is increased. Electron attachment to pentafluorobenzene appears to form at least two products arising from attachment followed by loss of an H atom and attachment followed by loss of a HF molecule. Work has begun on an analytical strategy to extend the range of molecules one can study with the Flowing Afterglow Langmuir Probe technique.

INVESTIGATIONS OF ELECTRON INTERACTIONS WITH MOLECULES: ELECTRON ATTACHMENT AND ELECTRON DETACHMENT REACTIONS OF HALOGENATED MOLECULES

Jane M. Van Doren   Albert D. Kowalak

Introduction

The interactions of electrons with molecules play important roles in many systems including electronics, biochemistry and atmospheric chemistry. One important area of research is the investigation of reactions which chemically transform electrons to negative ions. Such reactions play central roles in the production of ions in the atmosphere, determination of the atmospheric lifetime of neutral (reactant) molecules and the evolution of natural and man-made plasmas. The Air Force research program has contributed to our understanding of many of these processes. In recent years, the Air Force emphasis has been on plasma evolution and to a lessor extent, atmospheric lifetimes of trace species.

Information on interactions between electrons and molecules is relatively scarce, reflecting the difficulty in both experimentally probing electron-molecule interactions and interpreting the results of such probes. The Flowing Afterglow Langmuir Probe (FALP) Technique is a versatile technique which allows investigation of a wide range of electron-molecule reactions at low collision energies, providing quantitative values for electron attachment rate coefficients. Coupling the FALP with a mass spectrometer

allows product identification which is especially important for reactions in which more than one reaction

pathway is thermochemically allowed or in which the products formed change with temperature. In a small

number of reactions, two different product ions are formed. In this case, the mass spectrometer enables

one to determine the product branching fraction or relative rate coefficient for each product channel. The

experiments reported and discussed in this report were performed with the FALP with mass spectrometer

detection available at the Phillips Laboratory. This analytical research instrument is one of only three such

systems in the world.

The experiments discussed below focussed on reactions of electrons with a variety of

halogenated molecules. We chose to study these systems because data in the literature indicates that

introduction of halogen atoms into a molecule can result in the stabilization of the parent negative ion.

Species which can form stable anionic products upon addition of an electron can potentially do so

efficiently. Our investigations attempt to reveal some of the important chemical features which result in

efficient electron attachment and to provide information about the relative stability of the product ions.

Finally, in the process of our investigations, new applications of this experimental technique and new

analytical strategies were probed.

Experimental

The Flowing Afterglow Langmuir Probe Technique (FALP) utilizes a Langmuir probe which

measures electron density. In our experiment, a plasma is created by a microwave discharge in Helium (~1

torr). A small flow of Argon is introduced downstream of the discharge to eliminate the small fraction of

helium metastable species formed. The electrons and positive ions in the plasma are carried down the

length of the flow tube by a fast laminar flow of buffer gas. The electron density may be measured at any

point along the length of the reaction flow tube with a moveable Langmuir Probe. Neutral reactants are

added through a radial inlet downstream of the plasma. Ionic products of electron attachment reactions

may be identified with a quadrupole mass spectrometer which is located at the downstream end of the flow

tube and separated from the relatively high pressure tube by a small orifice. Most of the fast flowing gas is

pumped by a large roots blower and mechanical pumping system. A schematic of the instrument may be found in Figure 1.

A rate coefficient for electron attachment is determined from the relative decrease in electron density with reaction distance (which is proportional to reaction time in this "steady-state" experiment). The observed electron loss in the presence of an attaching gas is determined by diffusive loss to the flow tube walls and reactive loss from reaction with the reactant neutral. These loss rates are coupled. Therefore, loss by ambipolar diffusion is evaluated separately by monitoring the electron density as a function of distance in the absence of an attaching gas. Using this diffusion rate and the observed electron loss in the presence of an attaching gas, one numerically solves the equation as given by Smith and Spanel[1] for the electron attachment rate coefficient.

Results and Discussion

Most of the molecules investigated react with electrons by associative attachment, where the product ion is the anion of the reactant molecule. Other molecules were found to react with electrons by dissociative attachment, where the product ion is a fragment of the reactant molecule. Rate coefficients were determined and product ions identified for most of the reactions studied. The effect of temperature on the rate coefficient and product ions formed was evaluated for several systems. For associative detachment reactions, pressure dependences of the rate coefficient for attachment were evaluated. In some systems, subsequent detachment of electrons from the product ions formed was also observed. Preliminary analysis and discussion of this process is discussed below. Finally, several of the compounds investigated had either low vapor pressures or were highly reactive. These qualities make determination of electron attachment rate coefficients difficult. Following the lead of Smith et al.,[2] we have begun developing an analytical procedure for determining rate coefficients under these conditions.

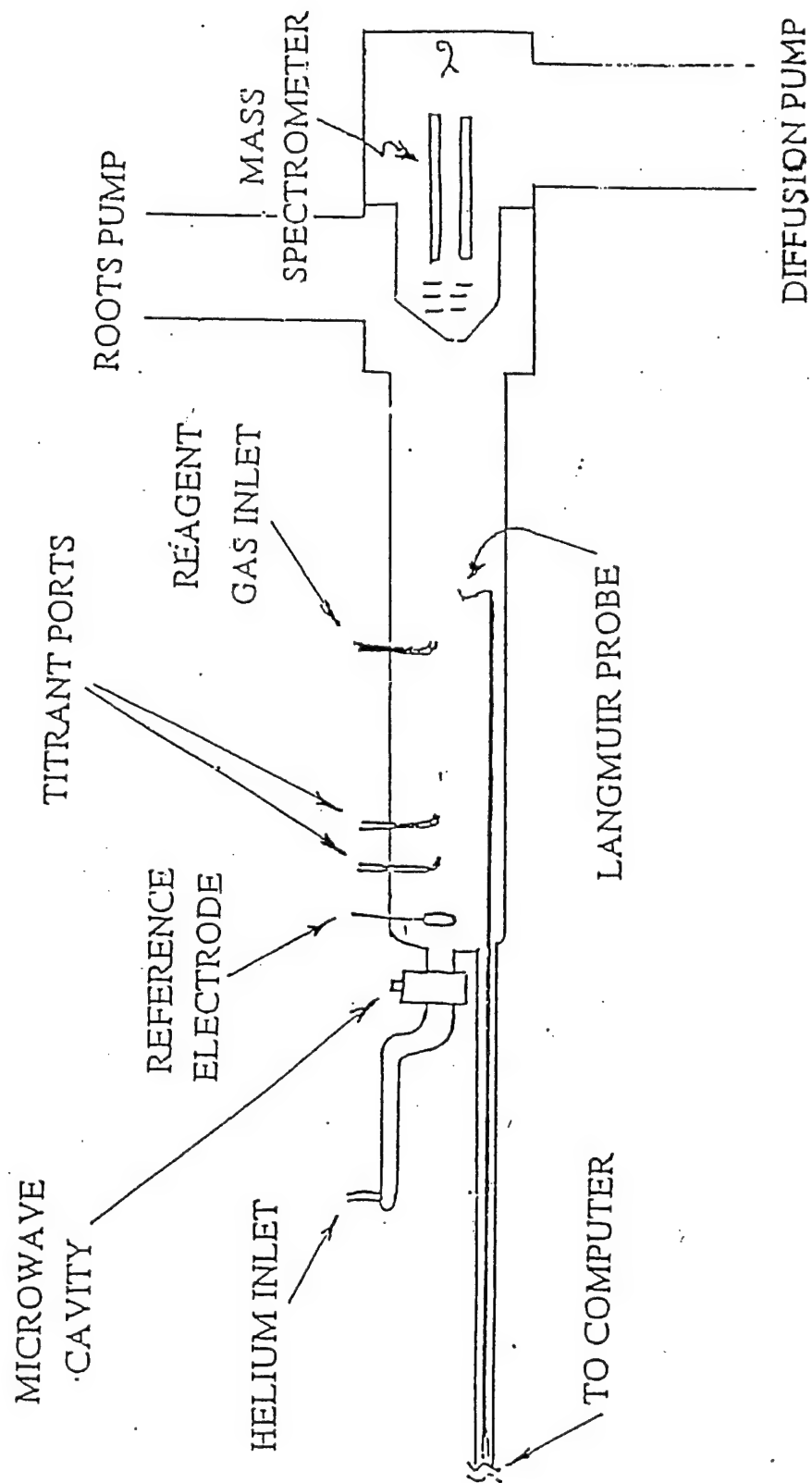# FLOWING AFTERGLOW LANGMUIR PROBE

Figure 1. Schematic diagram of Flowing Afterglow Langmuir Probe apparatus with mass spectrometer

Associative Attachment Reactions forming stable negative ions at 300K.

Octafluorotoluene, pentafluoropyridine, pentafluorobenzonitrile, pentafluoroacetophenone, trifluorotolunitrile all react with thermal electrons at 300K by associative attachment as evidenced by the mass spectra (see for example, Figure 2), i.e., only the parent negative ion is formed.

Electron attachment to octafluorotoluene and pentafluoropyridine was found to be very efficient with rate coefficients of 2.1×10^{-7} cm^3 s^{-1} and 2.1×10^{-7} cm^3 s^{-1}, respectively. Reaction of electrons with pentafluorobenzonitrile and pentafluoroacetophenone is also efficient. Accurate determination of the rate coefficient for electron attachment for these compounds proved difficult because the measured value appeared to depend on the length of time the reactant had been flowing in the inlet line (distinguished from the time for reaction with electrons in the flow tube). Such behavior, discussed further below , is indicative of species which can and do react with trace species adsorbed on the inlet walls or react with the inlet walls themselves.

An estimate of the electron attachment rate coefficient can be obtained for such systems by monitoring the observed rate coefficient as a function of time the vapor has been flowing in the inlet lines (see for example, Figure 3). In general the flow of the reactant is reduced from the value measured at the flowmeter by reactions with the inlet walls. At long times, reaction will be complete or will reach equilibrium and the neutral reactant flow will be equal to the value measured by the flowmeter. Therefore, the value obtained at long times should provide a lower limit to the true rate coefficient. This explanation appears to be supported by the data obtained for pentafluorobenzonitrile. Using the time dependence of the observed rate coefficient we estimate that the true electron attachment rate coefficient is $\geq 2.3 \times 10^{-7}$ cm^3 s^{-1}. In contrast, the measured rate coefficient for electron attachment to pentafluoroacetophenone did not appear to reach a maximum value over the time measured. At this point we are unsure of the implications of these measurements. We are currently developing an analytical procedure which would allow simultaneous neutral concentration and electron attachment rate coefficient determination (see below.) No rate coefficient was measured for electron attachment to trifluorotolunitrile because of the very low vapor pressure of this compound.
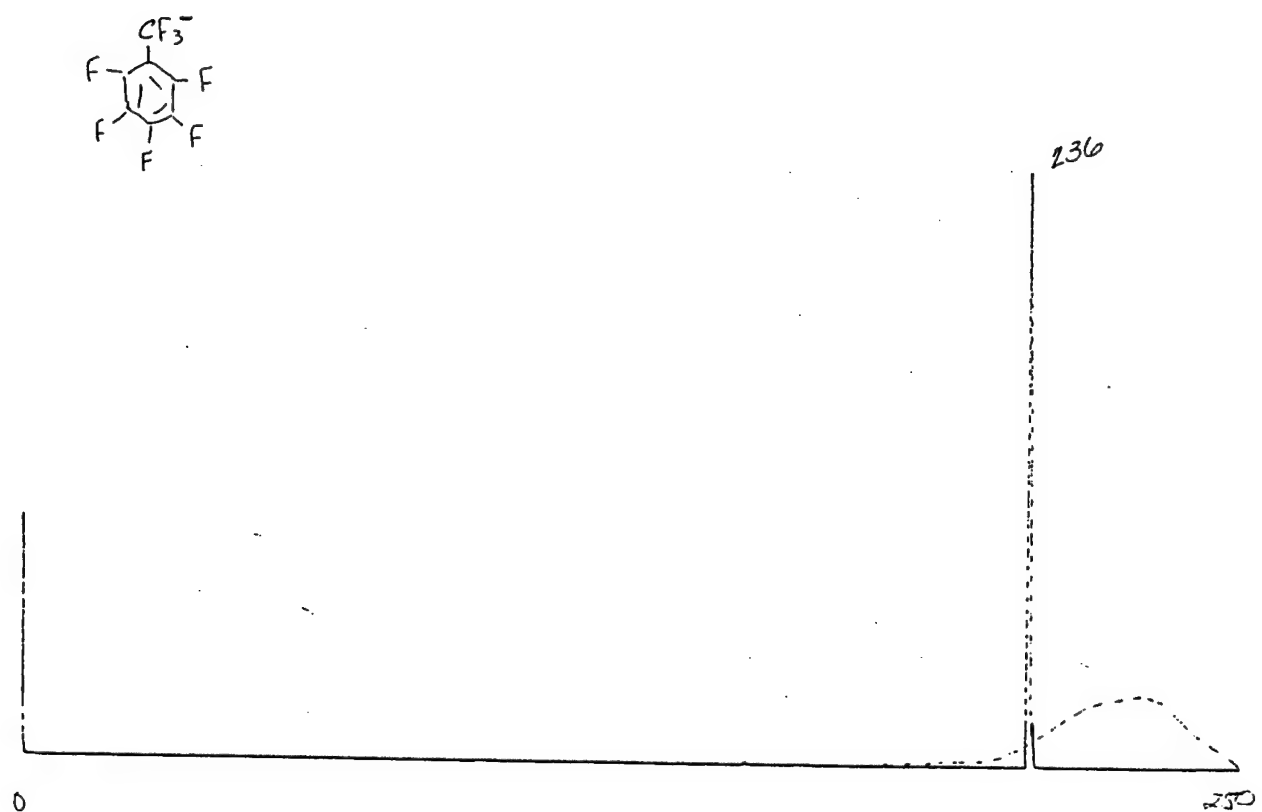
Figure 2. Mass spectrum obtained upon addition of octafluorotoluene to the reaction flow tube.

Figure 3. Observed electron attachment rate coefficient as a function of time the pentafluorobenzonitrile vapor has been flowing in the inlet lines .

Interpretation of the measured electron attachment rate coefficient for reactions forming an association product requires an understanding of the reaction mechanism. The initial association product is highly excited and has enough energy to detach the electron.

$$X + e^- \; \underset{k_d}{\overset{k_a}{\Leftrightarrow}} \; X^{-*}$$

$$X^{-*} + He \; \overset{\alpha k_c}{\Rightarrow} \; X^-$$

The bimolecular rate coefficient for association is given by k_a while the unimolecular rate coefficient for detachment of the electron from the energetically excited association product is given by k_d. Only those excited association products which undergo an inelastic collision with the surrounding molecules (typically He at pressure of 1 torr) to lose energy and form a stable (bound) negative ion result in reaction, or loss of electrons as measured in our millisecond time scale experiment. Assuming the the concentration of $[X^{-*}]$ is in a steady state, we find

$$d[e^-]/dt = - k_a[X][e] \{1 - k_d/(k_d + \alpha k_c[He])\}$$

In the high pressure limit, the rate of electron loss is independent of the pressure of collision gas He and depends only on the initial association rate coefficient. In this case, the measured rate coefficient would be equal to the association rate coefficient. We have measured the rate coefficient for electron attachment to octafluorotoluene and pentafluoropyridine as a function of flow tube (collision gas) pressure and found the observed attachment rate coefficients are independent of this parameter. Therefore, in the cases of octafluorotoluene and pentafluoropyridine the measured rate coefficients are a direct measure of the association rate coefficients. Similar studies must be performed for the reactions of the other associatively attaching species.

Associative Attachment Reactions followed by Electron Detachment

Weakly bound anions can detach in the reaction flow tube. When anions initially formed by attachment of electrons subsequently detach, the observed loss of electrons due to the presence of an

attaching gas decreases. In our steady-state flow reactor, this appears as a positive deviation from the electron density expected in the presence of a single well-behaved attaching gas. This deviation is most pronounced at longer distances because at those distances a substantial number of anions (which can detach) are present (see for example, Figure 4).

We observed evidence of electron detachment in the reactions of pentafluoropyridine (at temperature $T \geq 465K$) and octafluorotoluene ($T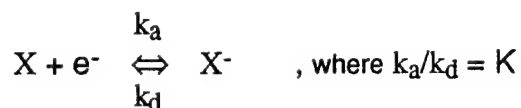 \geq 510K$). The only product ion observed in both reactions is the parent anion, the anion formed from associative attachment. In such cases, we can analyze the electron density data quantitatively to find values for the electron attachment and detachment rate coefficients. Specifically , we fit the observed loss of electrons with distance with values for the attachment and detachment rate coefficients assuming that electrons are lost by attachment and diffusion and "gained" by detachment. It is important to note that the constraints are provided by the stipulation that negative ions do not diffuse and that the sum of electrons plus negative ions must equal the number of positive ions at all times. As in our attachment rate coefficient analysis, the diffusion rate is not a variable but is determined in a separate experiment in which the attaching gas is not present. The recognition that these rate coefficients could be evaluated and the procedure discussed above was recently presented by Miller et al.[3] Electron attachment attachment and detachment rate coefficients for pentafluoropyridine and octafluorotoluene as a function of temperature are presented in Figures 5 and 6.

Interpretation of the observed rate coefficients requires an understanding of the mechanism of detachment. If the observed detachment is truly a thermal process the relative values of the rate coefficients, k_a/k_d, is a measure of the equilibrium constant K as defined by

$$X + e^- \underset{k_d}{\overset{k_a}{\Longleftrightarrow}} X^- \quad , \text{where } k_a/k_d = K$$

Such an analysis was presented by Miller et al.[3] This equilibrium constant is related to the exoergicity of electron attachment by : $\ln K = -\Delta G/RT$, or as a function of temperature, $d(\ln K)/d(1/T) = -\Delta H/R$. For electron attachment reactions the exothermicity of reaction is approximately equal to the electron affinity of the neutral molecule X. Therefore, by plotting the relative rate coefficients obtained as a function of

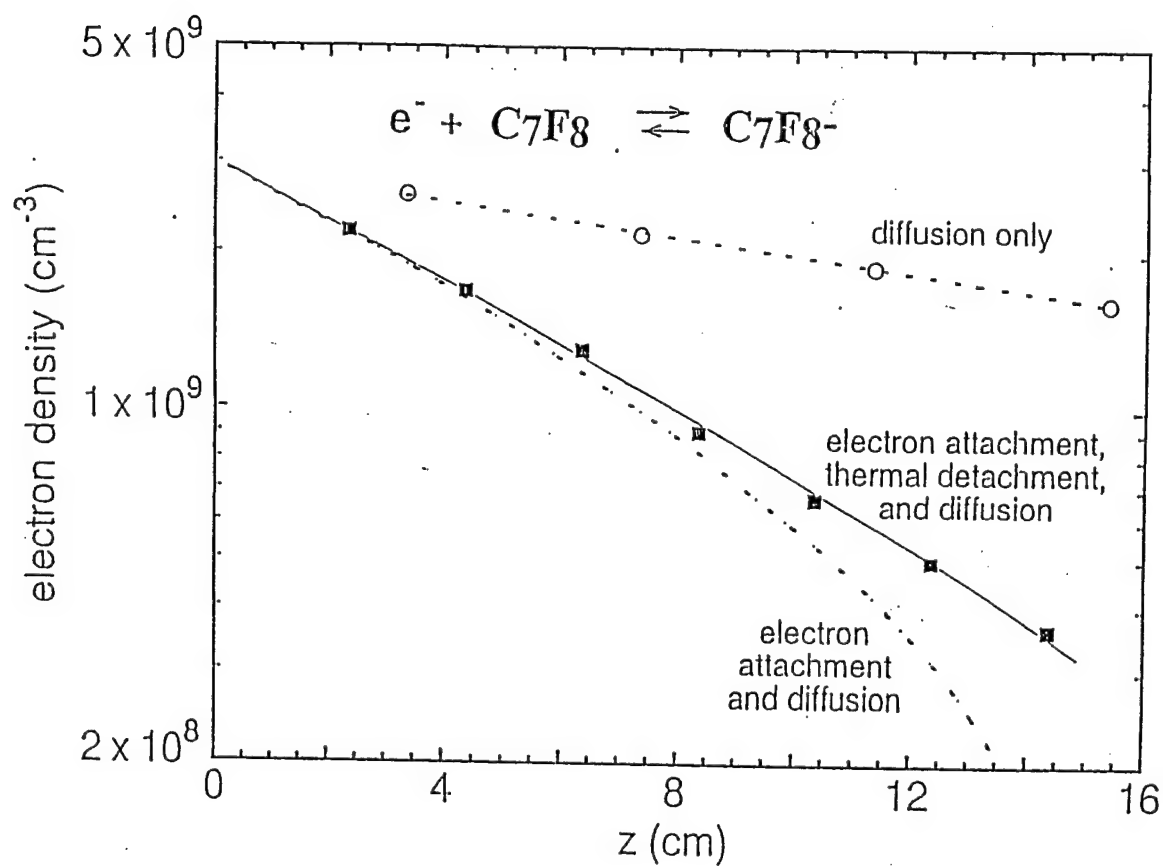ELECTRON DETACHMENT WAS OBSERVED AT
TEMPERATURES > 510 K

Figure 4. Electron density versus reaction distance when electrons are lost by diffusion only, lost by attachment to an attaching gas and by diffusion, and lost by attachment to an attaching gas and by diffusion but gained by detachment of negative ions.

1/T, a van't Hoff plot, we can determine the electron affinity. In principle the intercept of this plot should give the entropy of reaction. For our limited temperature range, this value will not be accurate. Analyzing our data for pentafluoropyridine and octafluorotoluene in this way, Figures 5 and 6, we obtained values of 0.83 eV and 0.72 eV, respectively. These values compare favorably with values estimated from the ion-molecule charge transfer measurements by Kebarle and coworkers (see Kebarle and Chowdhury[4]) who found values of 0.68 eV for pentafluoropyridine[4] and 0.83 eV for octafluorotoluene. [5] It is important to note that the relative electron affinity values for these two molecules obtained with our electron attachment-detachment data differ from those of Kebarle and coworkers by sign. We suggest that this difference arises from the lack of precision of our values; the experiment of Kebarle and coworkers directly monitors relative rates of electron transfer.

In order to determine whether this analysis is appropriate, we began exploring the implications of different microscopic mechanisms for electron detachment. As with associative attachment, such an interpretation as presented above, requires that the rate coefficients measured are pressure independent. Preliminary data on attachment and detachment to octafluorotoluene showed no pressure dependence. However, we found a pressure dependence for the rate coefficients for electron attachment to $c\text{-}C_4F_8$ and detachment from $c\text{-}C_4F_8$, a system previously investigated by Miller et al.[3] Interpretation of this pressure dependence was not straight forward. Extensive studies must be performed which quantitatively probe the effects of pressure, temperature, reactant neutral flow rate, electron temperature, and reaction time dependence.

Dissociative Attachment Reactions

Chloroacetonitrile and pentafluorobenzene react with thermal electrons at 300K dissociatively. Only one dissociative product channel is operative in the reaction of electrons with chloroacetonitrile while it appears that at least 2 product channels are operative in the reaction with pentafluorobenzene, i.e., two different product ions are formed in the reaction.

Chloroacetonitrile ($ClCH_2CN$) attaches electrons dissociatively to form Cl^- over the full temperature range studied (295-556K). While the thermochemistry of chloroacetonitrile is not well established,

Figure 5. (a) Electron attachment rate coefficient for reaction with pentafluoropyridine as a function of temperature (b) Electron detachment rate coefficient as a function of temperature (c) Equilibrium constant for electron attachment and detachment as defined in the text.
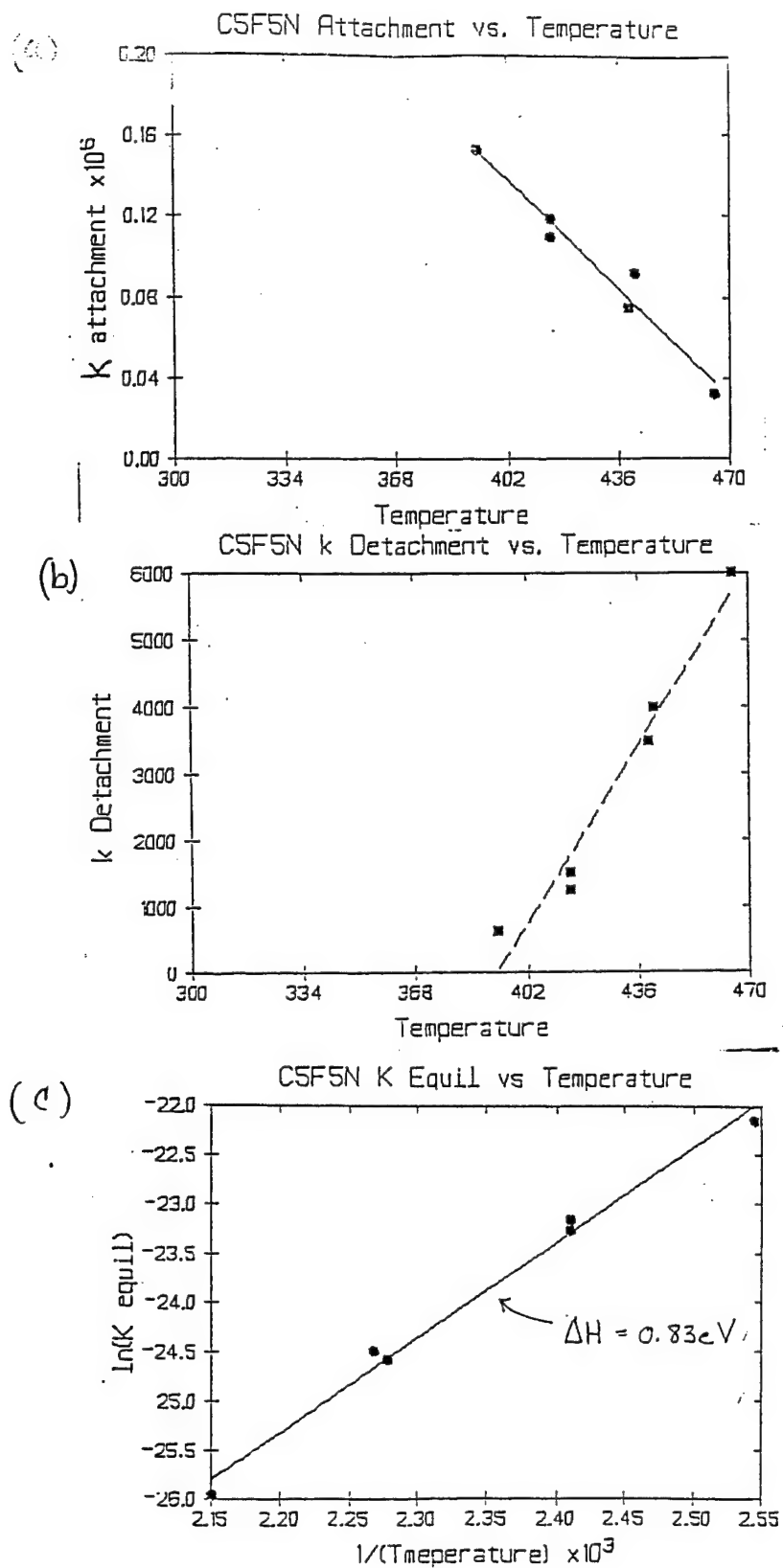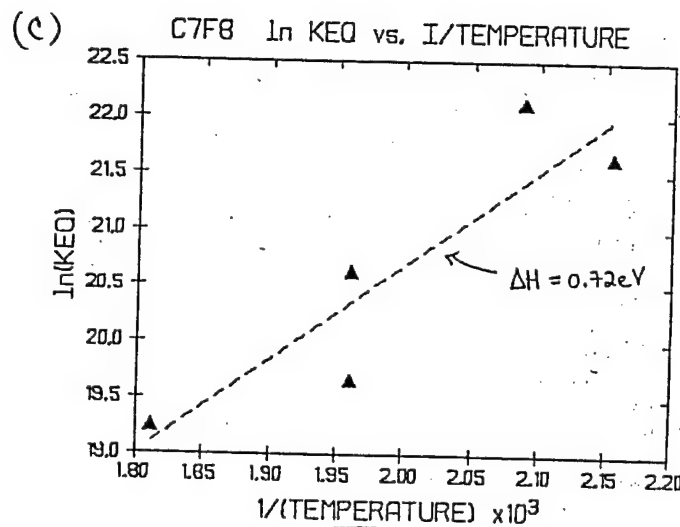
Figure 6. (a) Electron attachment rate coefficient for reaction with octafluorotoluene as a function of temperature (b) Electron detachment rate coefficient as a function of temperature (c) Equilibrium constant for electron attachment and detachment as defined in the text.

estimates[6] of the heat of formation of ClCH$_2$CN indicate that formation of Cl$^-$ + CH$_2$CN is the only

exothermic (and exoergic) dissociative reaction channel possible with ΔHrxn ~ -68 kJ mol^{-1}. No

information is available on the electron affinity of chloroacetonitrile so that we cannot assign an enthalpy

for associative attachment.

The rate coefficient for electron attachment to chloroacetonitrile is moderate at 295K, 4 x 10^{-8}

cm^3 s^{-1}, and increases rapidly to 1 x 10^{-7} cm^3 s^{-1} at 556K. Since the reaction is exoergic, we presume that

the positive temperature dependence arises from an activation barrier to dissociative attachment. A

simple Arrhenius analysis of this temperature dependence, gives a pre-exponential of 2.4 x 10-7 cm^3 s^{-1}

and an activation energy for this exothermic process of ~5 kJ mol^{-1}. In this model, the pre-exponential

factor can be considered a collision rate coefficient of sorts. Interestingly, this experimental "collision rate

coefficient" is approximately the value estimated using the Klots approximation for the true collision rate

coefficient. Despite the appeal of this simple analysis, it is probably not appropriate; the Arrhenius plot of

lnk vs. 1/T is not linear over our experimental temperature range. The problem with this simple analysis is

that it assumes a single rate limiting and temperature dependent process is occurring. Based on current

understanding for electron attachment reactions, we presume that the parent negative ion is formed

initially, independent of what product channel is observed, although the state of the parent negative ion

cannot be assigned. In other words, the mechanism for dissociative electron attachment to

chloroacetonitrile forming Cl$^-$ + CH$_2$CN can be described by:

$$e + ClCH_2CN + M \Rightarrow [ClCH_2CN^-]^* + M$$, where M is a collision partner (=He in our experiment)

$$\Downarrow$$

$$Cl^- + CH_2CN$$

The temperature dependence observed, therefore, combines the dependence of both the associative

electron attachment process and the subsequent unimolecular decomposition of the internally excited

negative ion.

Pentafluorobenzene reacts with electrons at 300K dissociatively. Major peaks in the mass

spectrum were at m/e = 98, 148, 167, and 184. Investigation of the peak height dependence on reactant

neutral flow rate indicates that the m/e 184 peak arises from secondary ion-molecule reactions of the primary products. The other three peaks arise principly from electron attachment. The anions with m/e = 167 and 148 correspond to attachment of an electron with loss of an H atom and a HF molecule respectively. Loss of HF upon addition of an electron is analogous to the reaction of benzene with O^- forming the benzyne radical anion.[7] We believe this is the first observation of the $C_6F_4^-$ anion. The anion formed from electron attachment followed by hydrogen loss has been observed in several other experiments but it's source was not pentafluorobenzene, as in our experiments. We do not currently have a hypothesis as to the structure of the anion observed with m/e 98. Based on our peak height versus flow rate analysis, the positive ion spectrum, and the structure of our reactant, we suggest that this anion is formed from an impurity in our pentafluorobenzene sample. Further work is needed to support the product identifications proposed and to quantify the electron attachment rate coefficient.

Other Reactions Studied

We investigated the reaction of electrons with hexafluoropropyleneoxide and ClF. However, uncertainties as to the purity of these species combined with the complicated mass spectra observed made interpretation of these data impossible. It is interesting to note that an anion was observed at the m/e corresponding to the parent ion of hexafluoropropyleneoxide which would indicate that the parent ion is bound, i.e., it has a positive electron affinity.

Development of an Experimental Procedure and Analytical Strategy for Highly Reactive and Low Vapor Pressure Compounds

As discussed above, it is often difficult to accurately monitor the flow rate of compounds into the flow tube. Both species which are highly reactive and those with low vapor pressures pose such problems. In the case of highly reactive species we have estimated rate coefficients by extrapolating the measured rate coefficients as a function of the length of time the reactant neutral vapor had been flowing in the inlet lines. Such a procedure not only takes experimental time and wastes resources; it is not a satisfactory solution. Following the lead of Smith et al.[2] we decided to probe the utility of using a

combination of second order kinetics and pseudo first order kinetics to solve for both the reactions rate coefficient and the neutral reactant flow rate.

The procedure involves measuring the electron loss with distance under conditions where the reactant neutral is in excess (pseudo first order kinetics) and under conditions were the reactant neutral and electron densities are similar (second order kinetics). These two experiments must use the same neutral flow rate, i.e., the electron density must be changed between experiments and the experiments must be performed on a time scale short, as compared with the drift in flow rate observed. Analysis of the data obtained should lead to accurate and unambiguous values for both the electron attachment rate coefficient and the reactant neutral flow rate.

Application of this procedure is complicated by the experimentally and theoretically limited ability to vary electron density over a wide range of values. Furthermore, analysis depends on the resolution of differences between the first and second order kinetics data and these differences are generally small. Finally, an accurate analysis critically depends on an accurate measure of the diffusion rate which can change significantly in the presence of a high mass attaching gas.

We have analyzed data for electron attachment to several of our reactive/low vapor pressure compounds but are not convinced of the uniqueness of our rate coefficient and neutral concentration values. For comparison, we obtained data under the prescribed conditions for well-behaved electron attaching reactants and are currently analyzing these data. We hope to better define experimental parameters which will lead to unique, accurate and precise determinations of the attachment rate coefficient and the neutral reactant flow.

Conclusions and Suggestions for Future Research

Halogenated aromatic molecules tend to form stable parent negative ions upon reaction with low energy electrons and many do so efficiently. Of the aromatic species studied, only pentafluorobenzene reacts with electrons dissociatively rather than associatively. It is likely that in this case, the associative product anion is not bound. Dissociative electron attachment was also observed in the reaction with chloroacetonitrile. Despite the exothermicity of this reaction, the efficiency for electron attachment

increases with increasing temperature suggesting that at least one aspect of the dissociative attachment process involves an activation barrier.

Weakly bound anions may undergo electron detachment. Electron detachment from the anion of octafluorotoluene (for $T \geq 510K$) and pentafluoropyridine ($T \geq 465K$) was observed. We have used the data obtained to quantify the relative rates of electron attachment and detachment. However, further work is needed to understand the fundamental meaning of these values.

Finally, development of an analytical method to facilitate accurate determinations for highly reactive species and low vapor pressure species has been initiated. Such a method will greatly increase the range of species which may be studied with this otherwise very versatile technique.

Future work will focus on completing comprehensive investigations of each reaction, including temperature and pressure dependence studies of the electron attachment rate coefficients. Extensive studies will probe the mechanism of electron detachment observed in our experiment. These studies will involve the systematic variation of pressure, temperature, reactant neutral flow rate, electron temperature, and reaction time dependence. To further our chemical insight into electron attachment reactions we will investigate electron attachment to molecules chemically similar to those investigated to date. These data will help to further reveal the influence of chemical structure on the electron attachment process. Finally, we will make both experimental design changes and procedural changes in order to facilitate the determination of electron detachment rate coefficients to highly reactive and low vapor pressure species.

Research Fellowships. One of the authors was also supported (for the second half of the summer) by the Batchelor Ford Summer Faculty Fellowship provided by the College of the Holy Cross. Both authors acknowledge and thank the AFOSR for financial support and RDL for administrative support.

References

[1] Smith, D. and Spanel, P., Adv. Atomic Molec. & Optical Phys., **1994**, 32, 307.

[2] Smith, D.; Spanel, P. Mark, T.D., Chem. Phys. Lett., **1993**, 213, 202.

[3] Miller, T. M.; Morris, R. A.; Miller, A. E. S.; Viggiano, A. A.; Paulson, J. F., Int. J. Mass Spectrom. Ion Proc., in press.

[4] Dillow, G. W.; Kebarle, P., J. Am. Chem. Soc., **1989**, 111, 5592.

[5] Kebarle, P.; Chowdhury, P., Chem. Rev., **1987**, 87, 513.

[6] Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G., J. Chem. Phys. Ref. Data, **1988**, 17, Supplement No. 1.

[7] Guo, Y.; Grabowski, J. J., J. Am. Chem. Soc., **1991**, 113, 1.

A HETEROGENEOUS PARALLEL ARCHITECTURE
FOR HIGH-SPEED IMAGE PROCESSING

Daniel W. Watson
watson@cs.usu.edu
Assistant Professor
Department of Computer Science

Utah State University
Logan, Utah 84322-4205

Final Report for:
Summer Faculty Research Program
Phillips Laboratory

August 1994

A HETEROGENEOUS PARALLEL ARCHITECTURE
FOR HIGH-SPEED IMAGE PROCESSING

Daniel W. Watson
watson@cs.usu.edu
Assistant Professor
Department of Computer Science
Utah State University
Logan, Utah 84322-4205

voice: 801-797-2440
fax: 801-797-3265

Abstract

Unlike single-processor computers, parallel processing systems have not "converged" on a single
architectural model. Many parallel architectures exist, representing different computational models, control
schemes, memory hierarchies, and interconnection topologies. A related and recurring problem for many
advanced architectures is that, for most of the applications that use them, peak performance is achieved only
a fraction of the time. These effects are due in part to the diverse computational needs between different
problem areas, and even within the same area. One way in which these effects can be attenuated is by
taking advantage of different computational resources available, and by carefully matching algorithm
segments to appropriate architectures. Heterogeneous computing is the intelligent use of diverse processing
hardware to meet distinct computational needs. In this study, heterogeneous computing concepts are
exploited in the development of a small parallel processing system designed to analyze image sequences.
The architecture developed is an aggressively miniaturized MIMD shared-memory parallel architecture
consisting of 12 TMS320C30 processors in a 3-D monolithic assembly of approximately two cubic inches.
Intended for image processing applications on space-based and flyable vehicles, the architecture has a
potential of 396 MFLOPS with frame rates of 200 frames/second at a resolution of 256x256 12-bit pixels.

A HETEROGENEOUS PARALLEL ARCHITECTURE
FOR HIGH-SPEED IMAGE PROCESSING

Daniel W. Watson

1. Introduction

Unlike single-processor computers, parallel processing systems have not "converged" on a single architectural model. Many parallel architectures exist, representing different computational models, control schemes, memory hierarchies, and interconnection topologies. A related and recurring problem for many advanced architectures is that, for most of the applications that use them, peak performance is achieved only a fraction of the time. These effects are due in part to the diverse computational needs between different problem areas, and even within the same area.

One way in which these effects can be attenuated is by taking advantage of different computational resources available, and by carefully matching algorithm segments to appropriate architectures. Heterogeneous computing is the intelligent use of diverse processing hardware to meet distinct computational needs. Heterogeneous systems exist at many levels, and can for example be composed of workstations on an existing local area network or by supercomputing facilities many hundreds of miles apart, connected by high-volume data links. In this study, heterogeneous computing concepts are exploited at the sub-architecture level in the development of a small parallel processing system designed to analyze image sequences. The architecture developed is an aggressively miniaturized MIMD shared-memory parallel architecture consisting of 12 TMS320C30 processors in a 3-D monolithic assembly of approximately two cubic inches.

Intended for image processing applications on space-based and flyable vehicles, the architecture has a potential of 396 MFLOPS with frame rates of 200 frames/second at a resolution of 256x256 12-bit pixels. Images are written directly to four multiplexed dual-port memory segments Memory is organized on an inverted pyramid basis: Large blocks of independent memory for image input and low-level processing are shared on a local region basis while only a small amount of global memory is shared by all nodes. Communication is supported with internode interrupts. A mutual exclusion mechanism is provided with logical atomic test/set flags. Hardware support is provided for barrier synchronization.

Heterogeneity is accomplished within the architecture in several ways. First computation of images is divided with respect to the available resources for each processor in the system. For example, those processors that have direct links to frame image memory are given responsibility for low-level image processing tasks that require manipulations of large amounts of image data. Conversely, higher-level operations that deal with less data-intensive, more cooperative intensive portions of the algorithm are relegated to those processors that have a faster path to global shared memory.

A second way in which heterogeneity is achieved is in the implementation of a distributed, hierarchical-domain memory system with both module and global shared memory to increase parallel access. In most image-processing application, low-level pixel operations require memory with high availability for image I/O but have limited sharing needs. High-level operations generally require less memory but do substantially more sharing. In the design developed in this study, system memory is partitioned into separate domains with differing levels of sharing. For example, image segments are written directly to 32K of dual-port memory in each of four independent modules. At the next level, 512K of memory is shared by three processors in each module. At the global level, 128K of memory is shared by all C30s.

A third way in which heterogeneity is exploited is the way in which the communication and synchronization hardware and software support mechanisms work together to provide different models of parallel computation. By implementing a mutual exclusion mechanism with logical atomic test/set flags, a shared-memory message-passing computational model cam be implemented. Additional hardware and software implements a barrier synchronization mechanism. In this way, a barrier-MIMD computational model can be implemented, which in turn can be used to support an SIMD-related computational model.

The report is divided into the following sections. In Section 2, background material on the fabrication process for the target architecture is presented. Section 3 provides background in heterogeneous processing and miniaturized image-processing architectures.   In Section 4, the architecture is described, and in Section 5, a development strategy used for the architecture is outlined.

2. Multi-Chip Module (MCM) Fabrication

2.1. MCM Background

The present architecture is the by-product of a comprehensive packaging feasibility study, which treats the aggressive miniaturization of various disparate electronic functions into a virtually monolithic assembly, hence "monolithic packaging." This section discusses the evolution of monolithic packaging and the considerations that led to the selection of the present architecture.

In certain classes of embedded military systems, such as interceptors and space systems, size and weight are critically managed resources. The advent of multi-chip module (MCM) technology [DoF93], a form of advanced microelectronics packaging, provides a means to eliminate the many, relatively bulky single chip packages that normally comprise a complex system. Even MCMs, however, only provide a partial solution, because they are often assembled in the same manner as the components they subsume: mounted onto printed circuit boards, which are placed into chassis assemblies, that are in turn connected together by cables and harnesses. Hence, while MCMs increase density, they may not always reduce size and weight, particularly if functionality and performance are increased.

Phillips Laboratory has investigated more efficient solutions through an approach referred to "monolithic packaging." By extending the efficiency of MCM technology at higher levels in the packaging hierarchy (with the highest level being the system itself), monolithic packaging has the potential to dramatically reduce the size and weight of electronic systems, while improving their performance.

The basis of monolithic packaging is a form of MCM technology known as High Density Interconnect (HDI) [Lyk93a], invented by the General Electric Company, originally in support of defense requirements. As a patterned overlay MCM approach, HDI achieves the highest density in a planar assembly. The patterned overlay fabrication sequence, shown in Figure 1, recesses components (usually integrated circuits) in a planarizing mechanical substrate. Thin dielectric films are then laminated onto the substrate and patterned, forming the various interconnections between components. This direct deposition method provides an extremely high electrical performance, while the very small feature sizes provide high wiring densities, lending to the fabrication of complex, high-speed electronic functions. The close placement allows planar component packaging efficiencies of greater than 80% compared to less than 20% for traditional single chip packaging approaches.
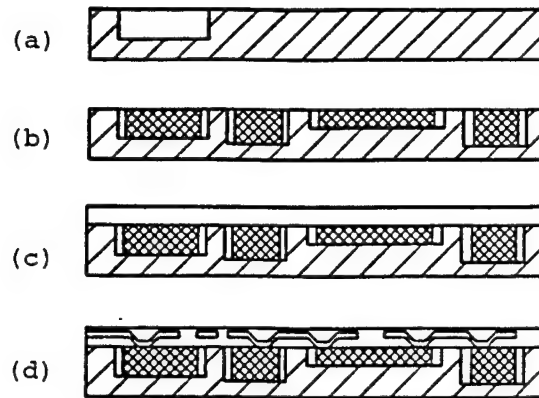
Figure 1: The High Density Interconnection (HDI) fabrication sequence. (a) Substrate milling. (b) Component placement. (c) Dielectric overlay. (d) Direct metal deposition.

The true benefits of monolithic packaging are realized by exploiting recently developed three-dimensional (3-D) extensions of the planar HDI process. By virtue of the planarity of both top and bottom surfaces of HDI substrates, it is very straightforward to form 3-D assemblies by stacking them together, as in a deck of cards. A simplified diagram of this process is shown in Figure 2. Individual substrates are affixed together with a special engineering thermoplastic that may be reworked, if required. An important key of this process is the formation of electrical contacts on the edge of the individual substrates. These contacts allow for level-to-level interconnections to be made through lamination and deposition processes similar to that used in the basic planar (2-D) HDI process.

Even though the 3-D HDI process is a monolithic packaging process, most work in this technology was based on low-profile stacked assemblies with a limited number of layers, most usually four substrates per assembly. Recent Phillips Laboratory research began to investigate the practical exploitation of a monolithic packaging approach that involves many substrates, or layers, per assembly. The technique of stacking electrical substrates (individual HDI modules) was conceptually expanded to include other classes of assembly, both electrical and mechanical (Figure 3). The monolithic packaging approach is similar to a micro-miniature, double sided backplane, except that the interconnections are formed directly and not all layers are electrical. With certain restrictions, a wide variety of functional and structural elements, ranging from heat-spreading plates for thermal management to entire inertial reference units, could be introduced into the monolithic MCM fabrication process. Through the use of engineering change or spacer layers, it is possible to introduce other MCM technologies and irregular, discrete electrical and mechanical components. As such, more comprehensive electronic functions could be realized, if not entire systems, through a single, monolithic assembly. Recent research demonstrations on 19-layer MCMs populated with interconnectivity structures have been performed, and thermal cycling test results on mechanical samples are encouraging.

38-6
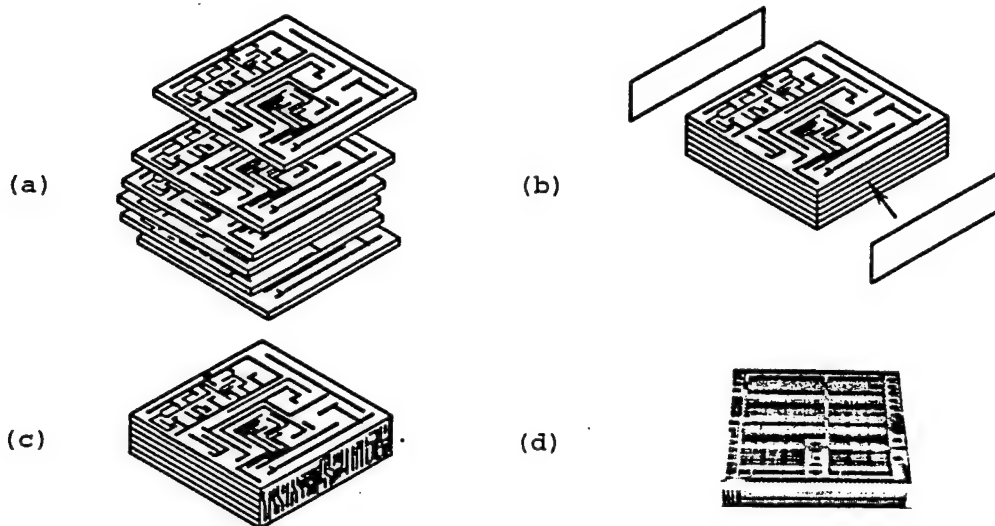
Figure 2:   Three-dimensional HDI process. (a) Substrate stacking. (b) Edge lamination. (c) Metal deposition and patterning. (d) Example of an operational 3D-HDI system, created for the NASA Goddard Space Flight Center.
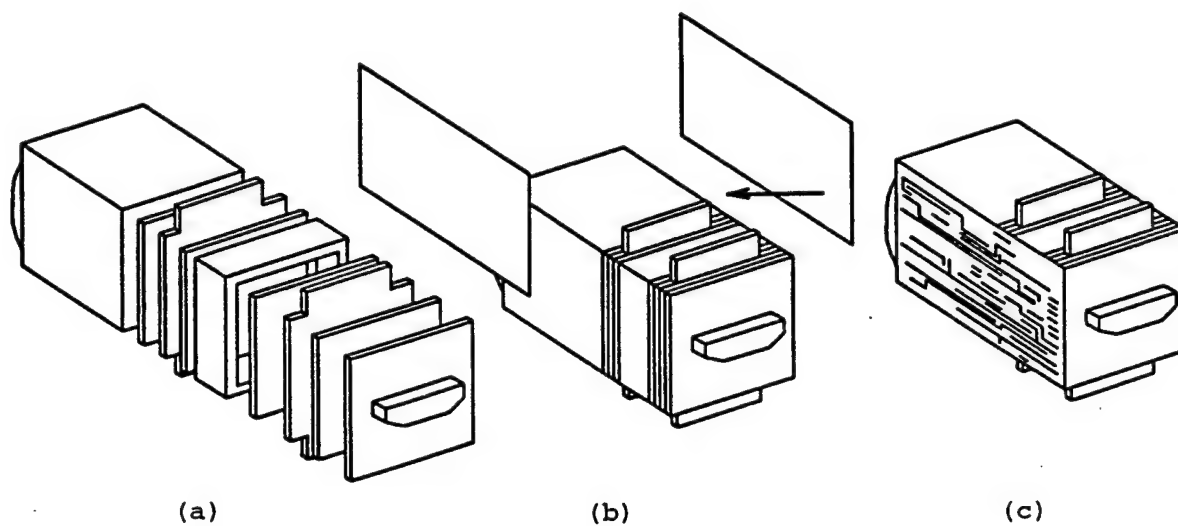
Figure 3:   Concept of a monolithic packaging process. (a) Exploded view, depicting HDI modules, imaging sensor front end, thermal plates, electrical spacer layers, power supply, and connector. (b) Assembled unit, before electrical overlays are formed. (c) Completed assembly.

38-7

2.2 Packaging Design Considerations for the Present Architecture

Packaging drives architectures, particularly in high-performance designs. Pin count, thermal management, and interconnection characteristics, for example, are important considerations that can affect the design of a computing system, whether packaged conventionally or monolithically. However, the most significant decision impacting the design of the architecture in this study is the selection of substrate size. In some respects, it may seem that choosing a large substrate size is most beneficial. Larger substrates provide more area for mounting components, and their larger perimeter accommodates a greater number of conductors for signals and power. Unfortunately, the fabrication of large substrates can be much more costly than small substrates for two reasons. First, the fabrication process used in MCMs operates on a "carrier lot" basis. A typical carrier contains five HDI MCMs, 1.6 inches square in size (the substrate size imposed on the present architecture) and two MCMs, 0.98 inches square. A 3.0 inch square substrate size, fabricated using the same carrier size, would yield only one MCM per carrier. The second problem associated with cost is the yield of fabrication. For the same reasons that semiconductor processors break apart wafers into dice, it is considered more practical to build smaller MCMs, because random defects will be less likely to occur on smaller substrates. Other mitigating factors further reinforced the decision to make smaller MCMs. In the case of an interceptor or sub-munition application, the small substrate size creates a more comfortable footprint.

2.3 Test and Assembly Considerations.

Significant to the current packaging research effort was the desire to explore new test and assembly approaches in support of monolithic packaging, for which the present architecture plays a key role. For test, new interconnection paradigms, such as the use of an interposer contact system, shown in Figure 4a, were desirable. Interposers are mechanically compliant electrical contact systems that are physically inserted between two planar surfaces which possess similar contact patterns. Previous General Electric research explored the use of "fuzz button" (Cinch Connectors, TRW) interposers, while more recent Phillips Laboratory research investigated mechanical interposers for hermetic applications [LyT92]. For monolithic packaging, the use of mechanical interposers on unpackaged HDI assemblies is desired. By using interposers, it is possible to test and operate 2-D MCMs prior to their combination into a 3-D, monolithic assembly. Using the interposer approach requires the formation of a matrix conductor pattern on the HDI assembly, resulting in the configuration shown in Figure 4b. The fanout pattern is the result of creating a one-to-one correspondence between the edge contacts for the 3-D overlay and the distributed contacts required by the interposer.
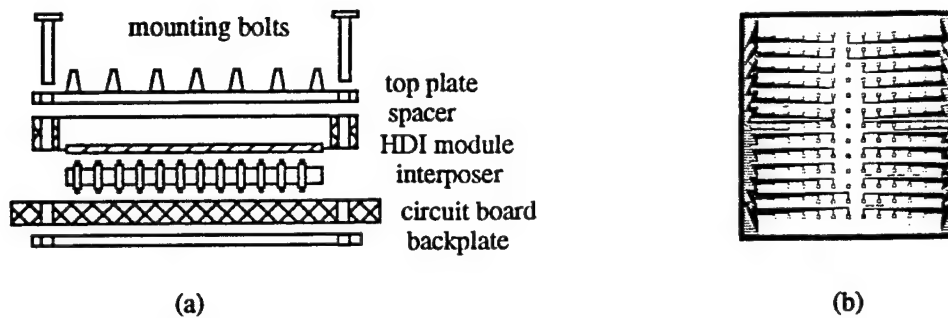
Figure 4: Interposer connection scheme for testing two-dimensional HDI modules. (a) Diagram of connection approach. (b) Example of HDI module, patterned for interposer connection, occluding the components underneath.

Conventional packaging approaches require normally permanent attachment of a substrate to a peripherally leaded package, which is then wirebonded. The interposer test approach, on the other hand, makes practical the construction of many-layered 3-D systems. After testing, HDI modules may be separated from a convenient fixture, inventoried, and aggregated for eventual fabrication as a 3-D assembly. The ability to pre-test individual modules enhances the chances of achieving a working 3-D assembly.

3. Architecture Background

One paragraph summaries of selected references:

[Bat82] - Usually regarded as on of the first works on airborne parallel processors. Ken Batcher's classic work where he follows up on his MPP and STARAN projects, and indicates that recent advances in VLSI technology make possible the development of an airborne parallel processing system (SIMD, distributed memory) using over 2000 bit-serial processing elements in a 0.5 cubic-foot box.

[BeA93] - Alliant Techsystems's Aladdin processor is an MIMD parallel processor with an SIMD co-processor with up to 256 processing nodes, all designed to fit in a soup-can sized package, perform target recognition using sensors of up to 2048x2048 pixels, and have a useful life of approximately 1000 seconds. They use 2-inch MCMs, and have a VME version available. C30s are the basic processor used. This is the most closely-related effort to the architecture under development as a result of this study.

[Fil93] - A taxonomy on next generation wafer technology - excellent reference for High-Density Interconnect (HDI). Topics discussed include thin-film MCM technologies, ongoing trends in semiconductor devices in general and MCMs in particular, and MCM materials and processes. GE's HDI is described in some detail.

[GiW92] - Good reference for processing images in a striped-subimage manner, rather than a square-subimage manner for several reasons including ease of implementation, and reduced communication overhead.

[MuB94] - Real-time motion detection using an active camera mounted on a pan/tilt platform. Use of morphological filtering is explored to desensitize the detection algorithm to inaccuracies in the background compensation.

[PaK90] - Mapping tasks of a multi-target tracking algorithm onto parallel computing architectures to maximize speedup. An asymptotically optimal (i.e., sub-optimal) mapping algorithm is developed and used to study effects of task granularity and architectures on speedup.

[RaS90] - Describes a special purpose co-processor for message passing. Similar problem to this work, but the authors took a different approach that appears to be more ambitious and general than our approach, which by contrast uses simple hardware changes to accomplish specific software needs to provide a more elegant, less development-intensive solution.

[SeW92] - AT&T's MCM DSP parallel processor. MIMD, 2-D grid, distributed memory, based on AT&T's DSP32C processor.

[YaS93] - Synthetic and inverse-synthetic aperture radar imaging for targets moving with unknown constant velocity in background. Good source for applications that can be targeted for the architecture in this study.

4. Parallel Processing Module Architecture

4.1 General

This section describes the processing module hardware along with its robust interconnection scheme. As shown in Figure 5, each processing module consists of:

- Three Texas Instruments TMS320C30 Digital Signal Processors
- One Xilinx XC3090 Field Programmable Gate Array (FPGA)
- 128k X 32 SRAM on a common shared primary bus (PBUS RAM)
- 16k X 16 bit dual-port SRAM on a common expansion bus (Image RAM)
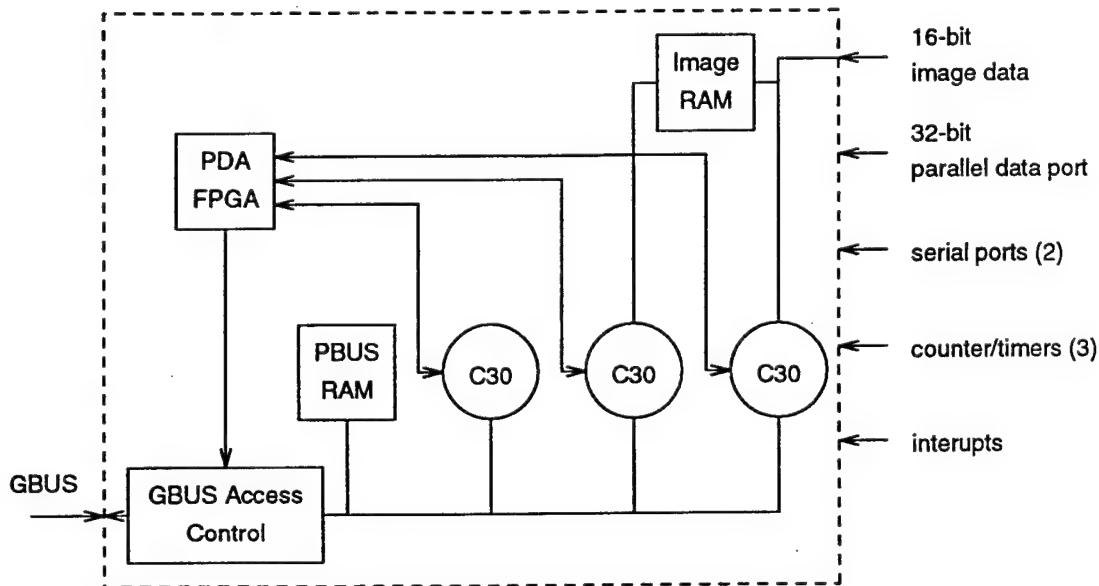
Figure 5:   Expanded processing module

By design, the data flow through the module begins at the dual-port RAM. From outside the processing module, the multiplexer module loads data directly into the image memory then sets a semaphore that interrupts the target DSPx. DSP1 & DSP2 cooperate to DMA and process portions of the image in common memory (again via semaphore). This processing compacts the raw data that is then placed in the PBUS RAM. It is further processed by the local master who places the further reduced data off-module to global memory. As a side note, the addressing scheme employed on our backplane allows the PBUS RAM to be accessed from other processing modules directly.

The design of the processing module stressed maximization of external I/O. Each credit card sized processing module has the following I/O:

- One A13:D32 parallel port
- One A14:D16 dual-port image memory port
- Two full-duplex 16.7 Mbaud serial ports
- Three programmable counter/timer pins
- Three processor direct interrupts
- Common backplane access, A15:D32 (VMEbus subset)

Although the original design of the modules was targeted specifically at meeting the real-time processing requirements of a single optical sensor, optimum I/O count on the module was attained to meet the needs of multiple sensor fusion applications. Each processing module can achieve 99.9 MFLOPS/50.1 MIPS of performance.

4.2 Intra-Module Architecture

The true capability of processing modules lies within the FPGA. Figure 6 shows the block diagram of the six functional segments within the FPGA. These functional blocks fall into four categories: bootup, PBUS control, interrupt distribution, and the external interface.
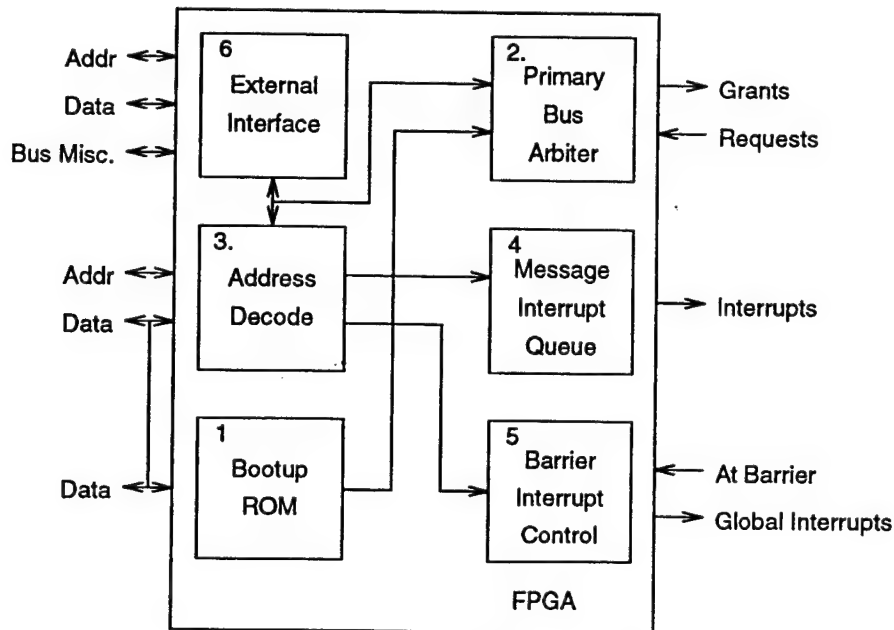
Figure 6: Functional Segments within processing module FPGA

The first FPGA segment functions only for processor initialization. At power-up, the FPGA enables a 64 X 32 bit segment from which the local master loads its bootstrap code. The bootstrap code configures the local master's serial ports to load applications to further configure itself and instructions to wake up the other C30s in an orderly fashion. A novel feature of this bootstrapping technique is that both the bootstrap code and, of course, the applications code is entirely reprogrammable because it is obtained from off-module. The bootstrap routine is truly a small set of instructions "firm-coded" into the FPGA simply to aim the bootup sequence to a serial port. The "firm-coding" of the bootstrap code is done by direct manipulation of the boolean state definition of specific CLB assets within the FPGA[Xil94].

The second functional segment is the PBUS Arbiter. In parallel architectures, processors require access to common blocks of memory. Many parallel architectures simply grant a fixed time-slice to all processors in sequence[Roc93]. This arbiter was designed to grant three types of PBUS access to those processors who request access.

As an example, consider only two of the three processors signal their desire to access the PBUS. A PBUS grant token of fixed length is passed between the requesting processors. No cycles are wasted polling the third processor because it had not requested PBUS access. When the third processor asserts its request for PBUS access, it is passed the bus grant. Once the processor is granted access to the PBUS, the address decode segment passes information to the arbiter segment about the duration of access.

Three access durations are allowed, short, long, and semi-infinite. These duration requests are important in that assure the processor of a fixed amount of cycles to access memory (e.g., test/set) before access is denied by the !HOLD. The net effect of this access methodology is that the programmer need only test a bit to see if it is its turn to access the PBUS. No other software actions are required, assuring the processors interleaved PBUS access. The third functional segment needs no discussion as its function is implicit by its name-address decoding.

The fourth and fifth segment allow for two types of intra-module interrupt handling. The first type is the barrier synchronization interrupt. Any processor writes a bit vector into a register within the FPGA. This bit vector represents those processors targeted to receive simultaneous interrupts. Through software magic, upon reaching a predefined code set point the given processor sets a bit that signifies it has reached a code barrier (i.e., "at barrier bit"). The instant the last processor, as described in the bit vector, reaches its set point an "everybody at barrier" interrupt is sent to all processors in that bit vector. Each processor clears the interrupt by de-asserting its "at barrier bit." This bit vector can be set or tested by any local processor during its PBUS access cycle.

38-13

The second type of interrupt handling is done by the fifth functional segment within the FPGA. This is the inter-DSP message handling interrupts. To interrupt a DSPx the requesting DSPy enters the PBUS, writes its message to local SRAM, then writes to the message interrupt address within the FPGA. During this write, at the assertion of !STRB, the target DSP is immediately interrupted. The interrupted DSP clears the interrupt by asserting its !IACK. Many message interrupts may be passed to a given DSP without it having clearing even the first one. This is due to the 16 message interrupt stack, per processor, using only one address location. The FPGA contains an addressable counter that drives the !INT1 line of each DSP directly.

The last functional segment within the FPGA, segment six, is the external interface. This section is designed as a subset of the VMEbus revision C.1 standard. More specifically, this interface behaves as an A15:D32 slave. Due to the low I/O count of our MCM package (i.e., 160 pins on a 1.6" X 1.6" substrate) the functionality available with the VMEbus was distilled. (Note: Just prior to the publishing of this study a higher density package has become available. This new package will be the carrier for subsequent design efforts.) Those selected portions of the interface maintain the straightforward addressing, bus request and interrupt handling found in VMEbus products.

4.3 Image Multiplexer Module

As with many focal plane arrays (FPA), the one used on MiP outputs several pixels that provide no useful information. The purpose of this module is to blank those useless pixels and direct the meaningful data to the dual-port RAM on each processing module.

Figure 7 shows the block diagram of the multiplexer module. The MiP FPA outputs 12-bit raw pixel data as a 268 X 256 matrix. The pixel blanker circuit operates by simply blocking the transfer of useless data through the module. In this design only a 256 X 256 array of data survives the pixel blanker circuit for direction to the processing modules. Each data frame is partitioned into subframes and targeted at 1-of-n processing modules by the subframe partitioner circuit. This channelization of data distribution yields MiP system throughput limited not by processor bandwidth but the analog signal processor(ASP) bandwidth! Since the ASP generates FPA data at 10 Msps, the MiP system processes this data at greater than 150 frames per second.
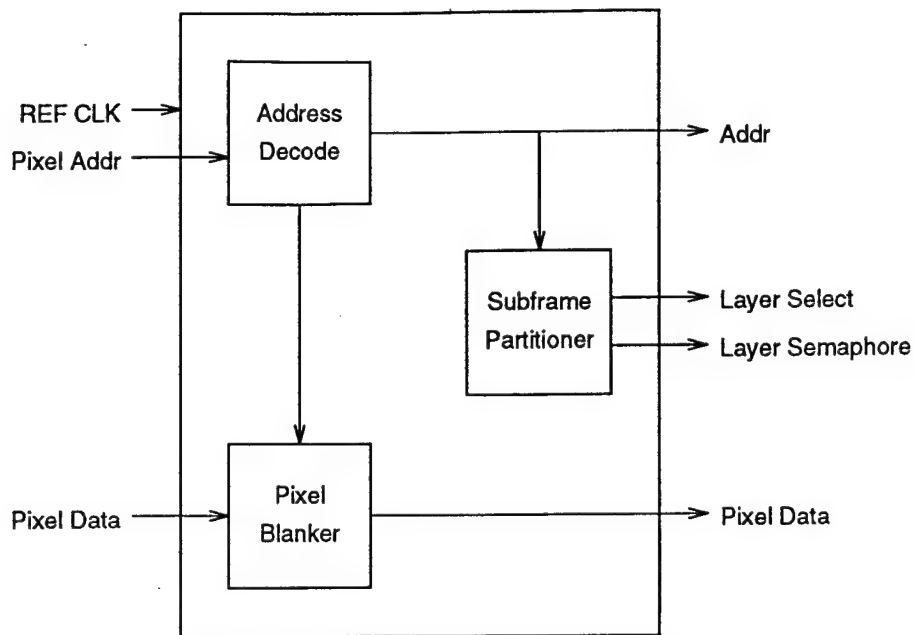
Figure 7:   Multiplexer Module

The Image Multiplexer will write the Focal plane image data to the image memories of each processing module on a round robin basis. The four image memories will each receive 1/4th of the total image data. For a 256x256 image, module 0 receives rows 0 through 63, module 1 receives rows 64-127, and so on. In this case, each module's image memory would be 16k.

When the multiplexer writes to one of the four image memories, it will set a corresponding bit in the Mux Status register. The Mux Status register will consist of four read-only bits representing each of the four image memories. The mux will set only one bit in the register at a time signaling that the memory is being written to by the Image Multiplexer and is therefore unavailable at this time. Regardless of the image data rate, the Mux Status register is zero 75% of the time for each processing module.

If a module is planning to read data from its image memory, it must first check the Mux Status register for a zero insuring that the Multiplexer is not currently writing to its image memory. If the register is zero when the processing module starts grabbing the image memory, it is possible that the Mux Status register will become a one before the module has finished reading all of the image data. In this case, the module is reading from the same memory block that the Image Mux is writing to. One unit must stop accessing the image memory. Since the Image Mux is in the process of writing an entire frame of data, It should not be halted. Therefore, the Mux will interrupt the processors when it starts writing to a new image memory block. The processors' interrupt handlers will then halt its reading process and start over by waiting for the Mux Status register to go to zero.

5. Design Motivation

To demonstrate the potential of encapsulating complex computing resources in an extremely reduced volume, a high-performance, parallel computing system was required, capable of performing sophisticated real-time processing for imaging sensors, communications functions, and platform control for a prototypical, next-generation interceptor. A prospective system concept revolved around a high-performance passive sensor (256 x 256 elements), an inertial reference unit, and various control and communications interfaces. Image data processing at frame rates of up to approximately 200 frames/second (14 mega-pixels/second) is needed. If possible, this concept system should also have the ability to interface with other sensors through parallel interfaces.

Given the state of the current monolithic packaging research, it was necessary to impose certain constraints on candidate architectures. With a 1.6 inch module and a conductor-to-conductor pitch of 18 milli-inches (mils), a standardized design of 80 conductors/edge on two edges was established for all substrates in the monolithic packaging system. The limitation of 160 conductors/substrate provided a challenge to candidate processing architectures. While these constraints were technology driven, the selection of the processor was ultimately to be based on the perception, real or perceived, that a radiation-hardened version would become available. This particular constraint was important to Phillips Laboratory, due to the focus on space systems, which necessitates consideration of the radiation environments that are inevitably encountered in space flight.

With the latter constraint, it was apparent that most candidate processor types were not oriented to real-time sensor processing. Few radiation hardened processors exist, fewer still with high performance. Fortunately, Phillips Laboratory has commissioned a conversion effort with Texas Instruments to establish a radiation-hardened version of the popular TMS320C30 digital signal processor (DSP) chip. It was in light of this consideration that made a parallel processing architecture from another packaging research effort quite attractive for the monolithic packaging research. This architecture, known as the Signal Processor Packaging Design (SPPD), was developed for Wright Laboratories (USAF) by Rockwell International (Anaheim, CA). Under the Wright Laboratories program, the SPPD had been built in brassboard and MCM form, but could not be used directly, because the form factor was substantially larger and would, therefore, not accommodate the monolithic packaging approach. Based on various configurations of multiple C30s, this architecture was selected for "simple repackaging" using the monolithic packaging fabrication, test, and assembly methods.

The SPPD architecture is based on multiple, identical bus-connected C30 networks (referred to as "P-busses"), connected to two global memory stores. As shown, each P-bus contains three C30s for the MCM configuration, while each P-bus in the brassboard configuration contains four C30s. In this version of the SPPD, local storage is provided for the C30s, whereas previous versions did not, except for the small internal RAM storage inherent in the C30. Along with the introduction of the SPPD architecture, Rockwell developed a simple operating system monitor and a useful complement of tools for the host system (VME-based Sun4/SPARC, running X-Windows) [Roc93].

After obtaining an SPPD brassboard on loan, courtesy of Rockwell, a new operating system was created to permit local study, since the original source code for the shipped monitor was not available. Simple applications were developed to demonstrate control and data acquisition with simulated imaging sensor data. Based on the Figure 10 architecture and the 1.6 inch MCM size, a partition scheme was chosen that placed an entire P-bus on an MCM, which could be repeated four times, with the global stores placed on separate MCMs. After extended acquaintance with the SPPD, it became clear that hardware modifications would be required for several reasons. First, the arbiter for the SPPD was not available, making it necessary to recreate this application specific integrated circuit (ASIC). Unfortunately, the information set provided, while extensive, was incomplete. Second, the pin count limitation of the monolithic packaging form factor chosen was inadequate to support two busses for global storage, and the size of the MCM did not permit a partition that would not have resulted in violating the 160 pin count limitation. Finally, the C30 input bandwidth limitation of 5 MHz (given a 33 MHz system clock) could not keep pace with the desired sensor data rate.

6. Summary and Future Work

The role of packaging in advanced architectures will become increasingly significant as performance continues to increase. To reduce propagation delay and avoid transmission line effects, it will be necessary to reduce the distance between components. This trend is difficult to maintain without the use of multi-chip packaging, because as pin counts increase, so does the area of a single chip package. The trend of pin count increase is consistent with Rent's rule and the trend for greater bandwidth. This trend also places demands on packaging technology. Two-dimensional advanced packaging technologies, such as MCMs can provide significant advantages for some architectures, if the highest performance elements are implemented within the MCM. Extended architectures, or architectures with many elements, will eventually consider three-dimensional packaging approaches for the same reasons. Monolithic packaging provides for the greatest performance efficiency of any currently defined approach, by virtue of component density, wiring density, and interconnection characteristics [Lyk93b].

The packaging technology underlying this architecture continues to improve. At the beginning of the present research project, a conductor-to-conductor edge pitch of 18 mils was established for individual MCM substrates, due to technology limitations that have since been overcome (it is now possible to practically achieve a 5 milli-inch pitch between conductors). If four edges had been used for the processor edge connections instead of two, an additional factor of two increase could be gained. Expansion of the pin count per MCM from 160 to 480 on two edges (or 960 for four edges) makes possible significant enhancements to the bandwidth and features of the present architecture, such as providing additional bus structures to improve system bandwidth/performance, and adding local memory stores for each processor.

References

[Bat82]   Batcher, K. E., "Bit-serial parallel processing systems," *IEEE Trans. Computers*, Vol. C-31, No. 5, May 1982, pp 377-384.

[BeA93]   Belt, R., Anthony, D., Brandt, J., Krantz, D., Lum, M., Shenoy, B., Vojta, M., Wehner, W., Pridmore, J., Pridgen, J., Black, E., and Daum, W., "The Aladdin Processor: a miniaturized target recognition processor with multi-GFLOP throughput," *Hardware and Forward-looking Infrared Issues in Automatic Target Recognition*, Vol. 1957, pp. 264-275, 1993.

[DoF93]   Doanne, Daryl A. and Frazon, Paul D., eds., *Multichip Module Technologies and Alternatives: the Basics*, Van Nostrand Reinhold, 1993.

[Fil93]   Fillion, R. A., "A forecast on the future of hybrid wafer scale integration technology," *IEEE Trans. Components, Hybrids, and Manufacturing Technology*, Vol. 16, No. 7, January 1993, pp. 615-625.

[GiW92]   Giolmas, N., Watson, D. W., Chelberg, D. M., and Siegel, H. J., "A parallel approach to hybrid range image segmentation," *Sixth Int'l Parallel Processing Symp.*, March 1992, pp. 334-342.

[Lyk93a]   Lyke, James C., et al., "Three Dimensional Patterned Overlay High Density Interconnect (HDI) Technology," *J. Microelectronic Systems Integration*, Vol. 1, No. 1, pp. 1-42, 1993.

[Lyk93b]   Lyke, James C., "Architectural Impact of Patterned Overlay Multi-chip Module Technologies," *SPIE Annual Conf.* (Orlando, FL), June 1993, pp. 285-296.

[LyT92]   Lyke, James C. and Tausch, Hans J., Jr., "Development of a High-Performance 800-Pin-Count Package for 3-D WSI Systems," *Government Microcircuit Applications Conf.*, November 1992, pp. 537-539,

[McP93]   McCarley, P., and Phipps, M., ``Signal Processor Packaging Design (SPPD)," *SPIE Architecture, Hardware, and Forward-Looking Infrared Issues in Automatic Target Recognition*, Vol. 1957, pp. 240-251, 1993.

[MuB94]   Murray, D., and Basu, A., ``Motion Tracking with an Active Camera," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 5, May 1994, pp. 449-459.

[PaK90]   Pattipati, K. R., Kurien, T., Lee, R.-T., Luh, P. B., "On mapping a tracking algorithm onto parallel processors," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 26, No. 5, September 1990, pp. 774-791.

[RaS90]   Ramachandran, U., Solomon, M., and Vernon, M. K,. "Hardware support for interprocess communication," *IEEE Trans. Parallel and Distributed Systems*, Vol. 1, No. 3, July 1990.

[Roc93]   Rockwell International, "Enhanced SPPD (Six-C30) Multichip Module Digital Signal Processor," *promotional pamphlet*, 1993.

[SeW92]   Segelken, J. M., Wu, L. J. Lau, M. Y., Tai, K. L., Shively, R. R., and Grau, T. G., "Ultra-Dense: an MCM-based 3-D digital signal processor," *IEEE Trans. Components, Hybrids, and Manufacturing Technology*, Vol. 15, No. 4, August 1992.

[Woj93]   Wojnarowski, R. J., et al, "Three dimensional hybrid wafer-scale integration using the GE high density interconnect technology," *Int'l Confe on Wafer Scale Integeration*, January 1993, pp. 309-317.

[YaS93]   Yang, H., and Soumekh, M., "Blind-velocity SAR/ISAR Imaging of a moving target in stationary background," IEEE Trans. on Image Processing, Vol. 2, No. 1, January 1993.

[Xil94]   Xilinx Incorporated, *The Programmable Logic Data Book*, publication PN0401224, 1994.

DETERMINATION OF SPACE DEBRIS FLUX
BASED ON A FINITE MASS

Wayne J. Zimmermann
Professor
Department of Mathematics & Computer Science

Texas Woman's University
P.O.Box 22865
Denton, TX 76204

DETERMINATION OF SPACE DEBRIS FLUX
BASED ON A FINITE MASS

Wayne J. Zimmermann
Professor
Department of Mathematics & Computer Science
Texas Woman's University

Abstract

This paper provides an overview of a problem related to space debris. It begins with a short review of the basic concepts and continues with a brief remark concerning existing models. Following this review we then present some models based on the principle that all debris consist of the material placed in orbit. Using the mass of the satellites, their associate components, the cross-sectional geometry of the selected altitude layer(s), two mass-dependent-population distributions and two altitude-dependent-population distributions the flux can be determined. The first model is based on the limited available mass, a single layer, a uniform mass distribution of the debris and the assumption that all debris is contained within the selected altitude bin. The next model removes the assumption imposed by the uniform mass distribution and replaces it with a distribution based upon current observations. It is then indicated that these models can be extended to multilayer models which can then be extended by applying two distinct altitude-dependent-population distributions. This approach defines the following four models: (1) n-layers, uniform mass distribution, uniform altitude distribution, (2) n-layers, uniform mass distribution, and an altitude distribution based on observations, (3) n-layers, an observational based mass distribution and a uniform altitude distribution, and lastly, (4) n-layers, observational based mass distribution and an observational based altitude distribution. These models ignore the inclination dependent population distribution. Inclusion of this distribution would generate four additional models.

DETERMINATION OF SPACE DEBRIS FLUX
BASED ON A FINITE MASS

Wayne J. Zimmermann

Introduction

From the beginning of the space age thousands of satellites have been placed into orbit by many nations. It is unfortunate that much of what is in orbit consist of inactive satellites, boosters and small parts produced during separation. This collection of manmade artifacts is a form of pollution capable of reducing the probability of system survival, if not making survival impossible. This manmade artifact can be partitioned into three broad categorie: payloads, boosters, and debris. Based on many observations much of it falls into the third category, debris.

Debris is the result of two events: (i) impact, the collision between two or more objects, and (ii) explosions. It is possible that with time the number of objects can increase to a point that a debris belt of sufficient density will encircle the earth making it practically impossible to maintain or orbit an operational satellite since such an environment would significantly reduce a satellite's lifespan. As a result, the continued increase of population poses a threat to all future operational satellites. To reduce the threat two problems are addressed. First there is the problem of mitigation which addresses two problems: (i) control of future debris generation and (ii) reduction the current levels of debris. The second problem is that of estimating the current levels of existing debris. The focus of the paper is to address an elementary aspect of this problem.

Debris poses a threat by it existance. A single particle can destroy a functional system. The question is: just how large must such an object be if it constitutes a threat? The current belief is that any debris having a diameter of 1 cm or larger would effectively disable a satellite hence the domain of interest is defined accordingly. The 1 cm is based on the velocity, mass, and material makeup of the debris. In his paper on the evolution of debris clouds, [1], Housen indicated that the greatest danger come from debris in the millimeter to centimeter size because of their size and density. Hence there is some debate concerning what size constitutes a threat. In both, 1mm or 1cm, cases the size of the debris is small enough to pose a threat while generally remaining invisible. Hence, assessment of the

population is currently not possible using current observational methods. As a results this leaves only those approaches based on computer simulation, theoretical physics, some laboratory experiments.

With regards to the reduction of existing debris little can be done. Of course several techniques are available for reducing newly generated debris. One technique is to reduce the amount of intentionally generated debris through the reengineering of the section separation procedures. Another reduction technique is to select orbits that provide for orbital decay of all component. Selection of such orbit would provide for the metication of much of the accidently generated debris. A third technique consist of venting the booster's fuel tanks of excess fuels thereby reducing the amount of accidently generated debris by explosions.

Concerning the problem of estimating the spacial distribution defined by the debris population several approaches have been taken. One method is to measure the environment which has been the focus of several measurement campaigns. The current methods used make use of either radar or optics. Both techniques have significant limitations. Radar has difficulty measuring size and observing nonmetalic objects while optics has difficulty in determining size and observing dull objects. Hence the acquired data needed to describe the LEO (Low Earth Orbit) environment has induced error. This error can be reduced by the application of appropriate statistical techniques assuming a correct selection of the type of distribution for the description of the debris population is employed. However the weakness in this approach lies both in the assumptions inheritent in statistical techniques employed and in the selected sampling techniques. As indicated the key assumption is the type of distribution. Clearly what is needed is a good sampling procedure but is this possible? Physical limitations imposed by the observational techniques constrain the observation time as seen when using optical procedures. Since optical techniques are based on measuring reflective light this approach requires that the observations be made in early dusk or dawn and are limited to approximately one hour. Hence a good sampling procedure is seldom available.

Why all the concern for determining the density of the debris or equivalently the flux level? One reason for concern is that the solution to this problem, flux determination of a trajectory tunnel, is important for the information it provides concerning probability of accidental impact during the life period. Such information plays a key role in the selection of a trajectory.

As indicated, in addition to the problem of sampling there is the problem of

observation. Currently, all debris having a actual diameter less than 10 cm is invisible. Further, there is the problem of determining the actual diameter of the debris regarless of the technique used in making the measurements. In both cases the diameter cannot be measured directly, hence it must be computed using either the albedo, in the case of optical measurements, or by using the RCS (radar cross-section) measurement acquired by radar. In both cases there is significant error. Hence statistical procedures must again be applied; procedures that must be validated experimentally.

An alternate to this empirical approach is that of modeling the phenomenon by using a combination of techniques. Included are various statistical estimations of some of the models parameters, deterministic calculations based on principles from physics.

To illustrate this marriage of techniques consider a problem related to a breakup. A flux level is dependent upon the number of objects in the region. Estimation of this value which requires that we estimate the number of newly created objects resulting from a breakup. Though the use of an enhanced model for the deformation of material we can acquire an estimate of this number. Hence molecular hydrodynamics codes based on the first principles are used to model explosions and hypervelocity impacts. The results of these modeling efforts are then used to describe breakup of an object which is used to create artificial debris database providing a distribution for the mass and velocity of the debris. In this way we can approximate the current state of the environment.

Remarks on Existing Models

Some models are purely empirical, hence they provide reasonable results for short-term forecasting. One such model is Kessler's Model, [2]. Based primarily on statistical evaluation of a data set the model has a number of limitations. One such limitation is its region of applicability. The model provide excellant result for whenever the altitude is less than 1000 km. For a complete assessment of Kessler's Model see [3].

Another model that is currently employed by the space debris research groups is referred to as EVOLVE. This is a model composed of several submodels. As indicated in the previous section, it makes use of a limited version of a breakup model by storing the element sets for a limited number of debris particles generated by the breakup. It also uses the element set associated with the historical data defined by the missions. The model then calculates the trajectories of each object and at specified time increments determines the flux. In addition the models generates artifical impacts. The number of such impacts is based

upon empirical data. The limitations of this model is defined by the limitation of each submodel and the number of objects it can maintain in its data base.

Single Layer Model

This model makes use of a single datum, the total mass of all objects placed in orbit. In addition the model assumes that everything is in orbit and that the debris is uniformly distributed in a single region. Figure 1 illustrates this situation.
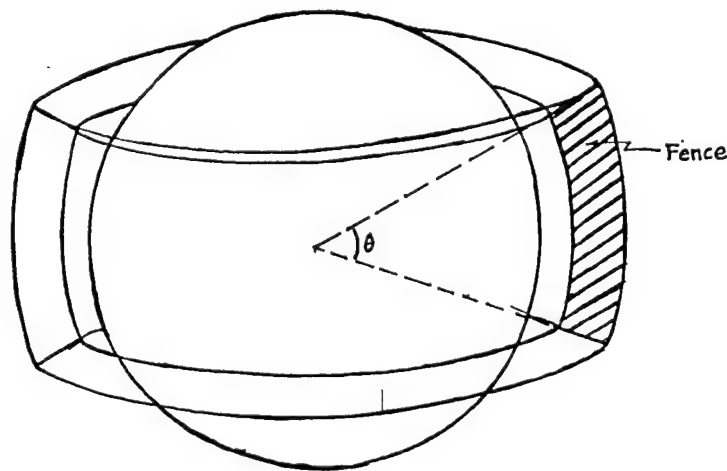
Figure 1. Single region environment.

As stated in Office of Technology Assessment paper, Orbiting Debris A Space Environmental Problem, [4], the mass of the 6500 space objects cataloged by the Space Surveillance Network (USSPACECOM) is about 2.0×10^6 kg. Using this single piece of data, the measurement of the earth's radius and geometrical relations defined by the trajectory tunnel[1] the flux can be determined given the assumption of the size and spacial distributions of the debris.

Three assumptions of the model are: the density of the resulting is somewhere between 1.00×10^3 and 2.72×10^3 kg, densities for water and aluminum respectively. A second assumption is: all resulting particles have the same density. Thoughout this paper density is taken to be 1.8×10^3 kg/m^3. Clearly the less dense the material the greater the number of particles since the particles are defined by size and the size of the original mass must be larger. Lastly, all are spherical.

To reiterate, this model is designed to provide an estimate for flux, an upper bound. Since most papers report the cumulative flux this paper will extimate the cumulative flux, that is, a measurement based on the number of all particles having a size greater than d_o. The cumulative flux will be denoted by F and it will depend upon a number of input parameters. Hence, F given by

$$F = F(\theta, d, \delta, m, h, \Delta, r_g, r_d). \qquad (1)$$

where

θ: denotes the width of the 'fence', which is measured in radian,

d: diameter of the debris,

δ: density of the material,

m_{tot}: total mass of all material put into orbit,

h: altitude,

Δ: tolerance for altitude band width,

r_g: rate of growth of mass,

r_d: rate of decay of mass.

The term fence denotes the cross-section of the toroid which contains all debris. Hence it is a two-dimensional region in a plane through the center of the earth and is contained in LEO.

From (1) it is seen that the model depends upon eight parameters. Throughout our discussion we will vary θ, δ, d, and h. The remaining parameters m, Δ, r_g and r_d will be assigned a value then they will remain fixed. As indicated the total mass $m = 2.0 \times 10^6$ and the average density $\delta = 1.80 \times 10^3$ kg/m^3, a weighted average of the density for aluminum, copper and plastic. A better approximation could be obtained by using a precise distribution of the various material used in constructing a satellite. But this might reduce the average density thereby increasing the flux. It will be seen, as noted earlier, that as density decreases the flux increases. For the rate of growth of mass and the rate of decay, r_g and r_d a respective, values of 0.04 and 0.01 are assigned and used throughout this paper.

With these remarks we now determine the flux based on the total mass of the space material, the density of the material, the size of the debris, assumed distributions for size and spatial density and the geometry of the region.

We now formulate the model. Let n_d denote the number of pieces of debris, m_d denote the mass of each debris particle, assuming each piece of debris is identical in size then

$$n_d * m_d = m_{tot} \qquad (2)$$

But the mass of each debris given by:

$$m_d = \rho * v$$

$$= \rho * (\pi\, d^3 / 6) \qquad (3)$$

The number of particles the total mass can create is denoted by n_d and is given as the ratio of the m_{tot}/m_d. Hence

$$n_d = 6\, m_{tot} / (\rho\, \pi\, d^3) \qquad (4)$$

Based on the definition of a fence, which is the cross-sectional area of a user selected toroid as illustrated in Figure 1, and the width is defined by the angle created by the rays extended from the center of the earth through two latitudes, the area of the fence is given by:

$$A_f = (\theta / 360)\, \pi\, (r_o^2 - r_i^2) \qquad (5)$$

where $r_o = R_e + h + \Delta$ and $r_i = R_e + h - \Delta$.

Assuming a circular orbit the velocity is given by:

$$V_e = \sqrt{(\mu / r)} \qquad (6)$$

where r denotes the radius of the great circle in the center of the toroid, that is, $r = R_e + h$. Using the circumference of the geometrical center of the toroid as the average distance a typical debris particle has a period of:

$$T = (\mu / r) / (2 * \pi * r) \qquad (7)$$

where the velocity is given in km/hour and μ is defined to be 3.986×10^5.

Recalling that flux is defined as the number of impacts per meter square per year we have

$$F = n_{rev} * n_d / A_f \qquad (8)$$

where n_{rev} denotes the number of revolutions per year and is defined by

$$n_{rev} = 8760 / T \qquad (9)$$

Recall there are 8760 hours in a year. Hence (8) can be written as:

$$F = K \, m_{tot} / [(Re+h)^{2.5} \, d^3 \; \theta \; \Delta \; \delta] \qquad (10)$$

where R_e=6378km, the radius of the earth, and

$$K = 3600*24*365*6*90*sqrt(\mu)/(2\pi^3) \;\; = 1.7338 \times 10^{11}$$

Table 1 below defines the parameters selected in determining the flux. The altitude is taken to be 700 km because of the recent interest in that region.

Table 1. Selected parameters for mass, average density, growth rate, decay rate, bandwidth time and altitude.

m_{tot} = total mass = 2.0×10^6
δ = density = 1.8×10^{12}
r_g = rate of growth = 0.04
r_d = rate of decay = 0.005
Δ = half band thickness = 200 km
t (years) = 1990
h = altitude = 700 km

Based on the above model and the parameters definition given in Table 1, Table 2 presents flux as a function of fence "width" and debris size.

39-9

Table 2. Flux as a function of fence width and debris size.

dia\width	40°	60°	80°
1.0	1.1×10^{1}	7.3×10^{0}	5.7×10^{0}
10.0	1.1×10^{-2}	7.3×10^{-3}	5.7×10^{-3}
20.0	1.4×10^{-3}	9.5×10^{-4}	7.1×10^{-4}
50.0	9.1×10^{-5}	6.1×10^{-5}	4.6×10^{-5}

Single Layer with Mass Distribution

In the previous model the total mass was partitioned into n_d debris particles of equal size. The intent of this section to embed the mass distribution into the model. This is done by using the weights in Table 3 to calculate the mass allocated to each size class. Clearly this additional information will reduce the flux level since the number of debris particles created is reduced; some of the mass goes into the creation of larger debris particles. This results in the reduction of the number of impacts and hence the flux.

The mass distribution is given by two distinct functions based on the process by which the debris is generated. The two processes which generate debris are, collision and explosion, create distinct distributions.

The distribution used to describe the number of fragments of mass m or greater for the collision process is:

$$N(m) = \alpha \ (m/m_{tot})^{-\beta} \tag{11}$$

where α and β are empirically determined constants which are depend on the impact conditions, m_{tot} the total mass and m the fragment's mass. Yates and Jones, [6], reported typical values of are 0.4478 and 0.7496 for α and β respectively.

Referred to as the power law the above distribution is scale invariance and has numerous applications in the theory of fractiles. For our discussion we will restrict our investigation to this distribution.

Since debris is typically specified by radius and we have assumed a fixed density then

the power law takes the form:

$$N(r) = \alpha \ (r/r_{tot})^{-3} \ \beta \qquad\qquad (12)$$

Based on this size dependent formulation of the power law we generate the values stated in the following table.

Table 3. Mass (Size) Distribution--Size is specified since the density
is taken to be 1.8×10^3 kg/m^3.

Size (cm)	Frequency	Percent
1	4127	0.792
2	657	0.126
3	210	0.040
4	91	0.017
5	47	0.009
6	27	0.005
7	17	0.003
8	11	0.002
9	7	0.001
10	5	0.001
11	4	0.001
12	3	0.001
13	2	0.000
14	1	0.000
15	1	0.000
16	1	0.000
17	1	0.000

To calcuate the number of particles the percentage is used to allocate a corresponding mass to the related size bin. Using equation (12) and the diameter defining the size bin we calculate the flux associated with the selected bin which is referred to as the bin-flux. The bin-flux is also calculated for all sizes greater than the selected size by summing over each bin-flux greater than the selected size thereby yielding the cummulative flux. Table 4

presents a representative output given the parameters stated in Table 1.

Table 4. Flux as a function of fence width and debris size when the
power law mass distribution is applied.

d/width	40°	60°	80°
1.0	2.9×10^0	1.9×10^0	1.5×10^0
10.0	2.7×10^{-4}	1.8×10^{-4}	1.4×10^{-4}
20.0	1.9×10^{-5}	1.3×10^{-5}	9.3×10^{-6}
50.0	5.8×10^{-7}	3.5×10^{-7}	2.6×10^{-7}

Although not included in this study, the situtation involving an explosion is discribed by the dstribution:

$$N(m) = (M/2\eta^2)e^{-\text{sqrt}(m/\eta 2)} \qquad (13)$$

where η is a Mott constant, see [7]. Using the (13) we can weight the debris accordingly. This is beyond the scope of this paper.

Conclusions

While the approach taken in this paper places an upper limit on the flux, it does so at the expense of assuming that the debris is uniformily distributed. Debris need not be uniformily distributed. If a large class of satellites are placed in orbit with a given inclination and partitioned into several planes then the points at which the orbits intersect will show a higher flux level. Something similar to the intersection of several roads.

It suffices to note that these models can be extended to multilayer models, which can then extended by applying two distinct altitude dependent population disttributions which are defined by the following four models: (1) n-layers, uniform mass distribution, uniform altitude distribution, (2) n-layers, uniform mass distribution, and an altitude distribution based on observations, (3) n-layers, an observational based mass distribution and a uniform altitude distribution, and lastly, (4) n-layers, observational based mass distribution and an observational based altitude distributon. All of these models ignore the population

distribution which is dependent upon the inclination. Inclusion of such a distribution would permit for the creation of four additional models.

References

[1] Housen, K.R., "The Short-Term Evolution of Orbital Debris Clouds", Jour Astronautical Science, Vol 40., No.2, 1992 pp203-213

[2] Kessler, D.J., Reynolds,R.C., Anz-Meador,P.D., "Orbital Debris Environ-ment for Spacecraft Designed to Operate in Low Earth Orbit", NASA Technical Memo 100 471, NASA Lyndon B. Johnson Space Center, Houston, Tx.

[3] U.S.Congress, Office of Technology Assessment, "Orbiting Debris: A Space Environmental Problem--Background Paper", Washington, DC, US Government Printing Office, OTA-BP-ISC-72, Sept., 1990

[4] Callaham, M., Carnett,C., McKnight,D., "Orbiting Debris-A Space Environmental Problem", Congress of The United States Office of Tech. Assessment Washington, D.C., OTA-BP-ISC-72

[5] Chobotov, V.A., Orbital Mechanics, AIAA , Washington, DC, 1991

[6] Yate,K.W., Jones,F.M., Assessment of the NASA Evolve Long-Term Orbital Debris Evolution Model, Phillips Laboratory Report, KAFB, NM., 1993

[7] Langton,L.W.,Ed. Textbook of Ballistics and Gunnery: Vol I, Her Majesty's Stationery Office, London, 1987